$t\bar{t}\rightarrow\tau$+jets Analysis
Pre-Blessing talk

Giorgio Cortiana,
University of Padova and INFN

CDF-7689

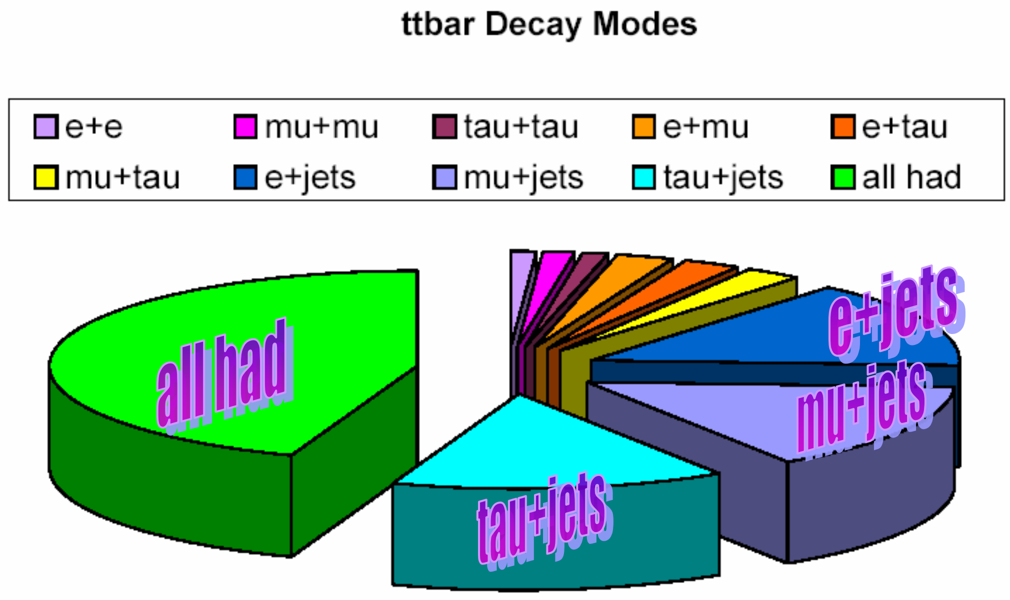
Main changes w.r.t previous results:
Lepton-id Scale Factors application

Outline:

- *Introduction and motivation*
- *Datasets and method summary*
- *Method-I background prediction*
- *Kinematical selection optimization*
- *Preliminary results and cross-checks*

Introduction

tau+jets



- ☺ **Not observed yet, can add a new piece of knowledge in the top physics sector**
- ☹ **Large background: QCD, EWK+HF**
- ❑ **taus are reconstructed as jet-like objects: more challenging signature compared to e/μ.**
  - ❑ **Will use MET rather than lepton ID: Extra Acceptance from “dirty” e/μ+jets events**

In order to successfully finalize a cross section measurement we

- **need an optimized kinematical + topology selection**
- **need b-jet identification to increase S/N ratio (SecVtx tags).**

- **b-jet identification rates are different on ttbar and background processes: can distinguish the two components:**

  → **B-tag rate parametrizations** → **Tagging matrix**

Datasets & Method

Good Run List: *v7.0 (1,1,4,1)*

Datasets and trigger:

TOP_MULTI_JET ***dataset*** **(gset0d) up to Aug 2004: 311 pb^{-1}.**

- **L1: ≥1 cal. tower with E_T≥10 GeV;**
- **L2: ≥4 cal. clusters with E_T ≥ 15 GeV, ΣE_T ≥ 125 GeV;**
- **L3: ≥4 jets, R=0.4, E_T ≥10 GeV**

MC : (167 fb^{-1}), Pythia ttbar (ttopel), M_{top} = 178 GeV

Offline version: ***5.3.3_nt5***

Jet Corrs: ***jetCorr04b***

High missing Et Selection?

yes → **τ+jets analysis**
ttbar→blν bbarjj

no → **All-Had analysis**
ttbar→bjj bbarjj

Method-I approach
+
ad hoc Kinematical selection

- **Method 1: positive tagging matrix approach to predict the absolute amount of background**
- **Kinematical Selection + ≥ 1 SECVTX positive tag**

ttbar cross section measurements in multi-jet final states

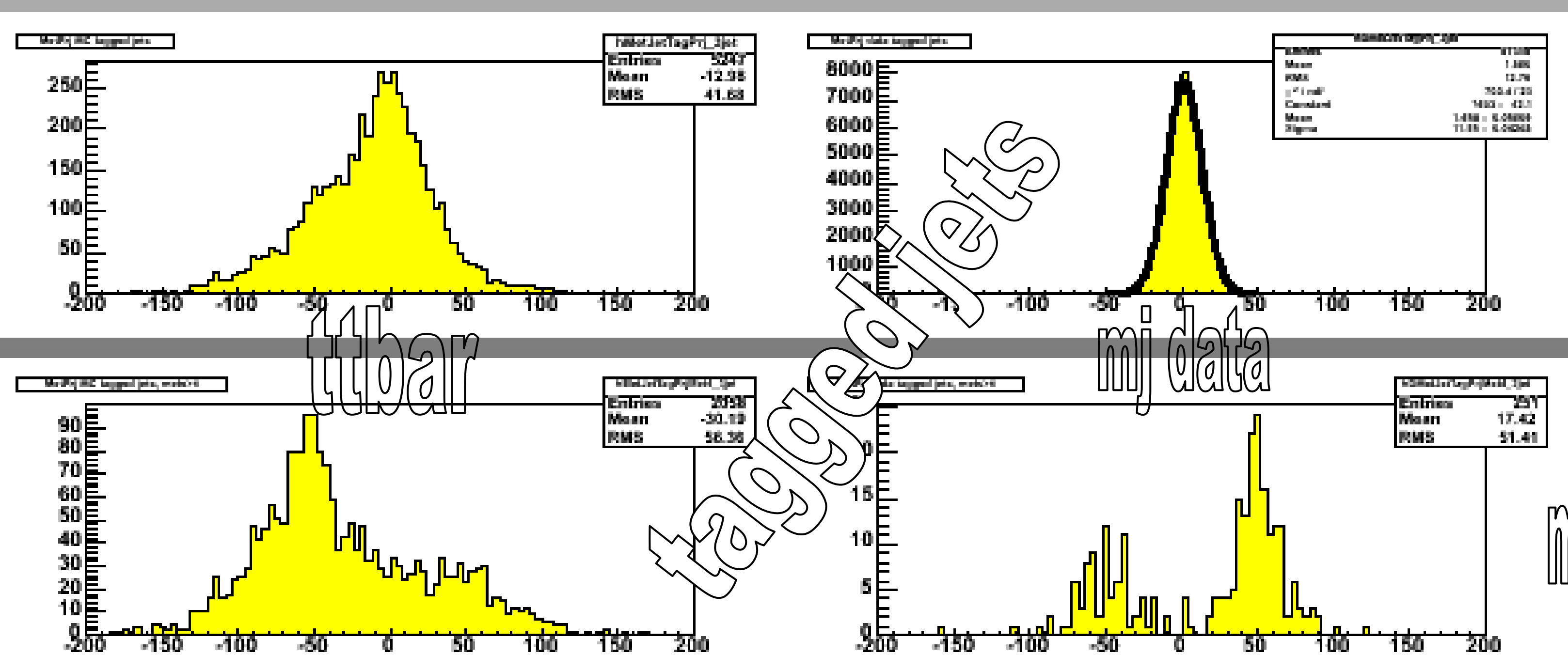
Method-I Background Prediction

- b-jet identification rates are different on ttbar and background processes: can distinguish the two components:

N evts	3 jets	4 jets
MJ no kin sel	879,187	1,553,525
MC no kin sel	16.88	182.92
MJ (mets≥4)	2,317	2,412
MC (mets≥4)	6.54	56.92

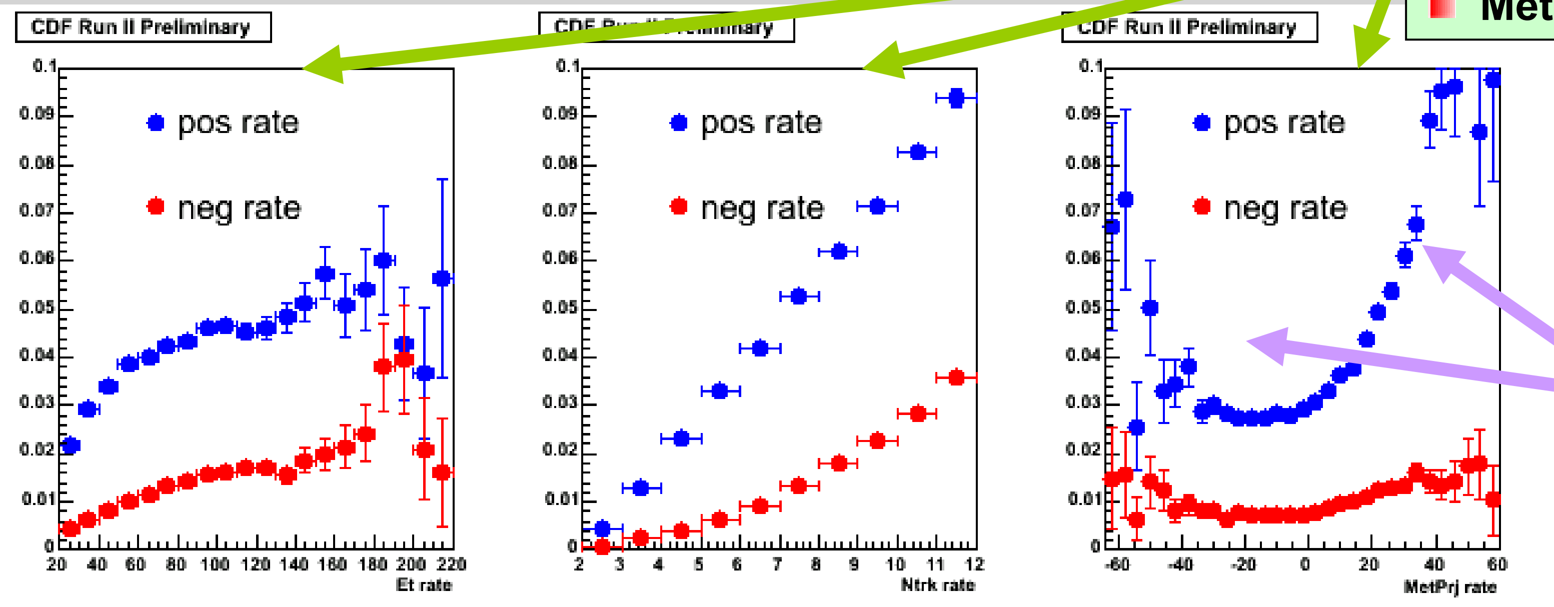
- Look at the B-tag rates directly from TOP_MULTI_JET data
- Use 3 (E_T^{L5} > 15 GeV, $|\eta|<2.0$) jet events: $F_{top} = 2\times10^{-5}$
- Take the vars by which the tag-rate mainly depends to construct a matrix

Met *cos Δϕ(met,jet): has a consistent correlation with heavy flavor component of the sample and allows to distinguish met origins in relation to geometrical properties



Method-I background prediction - 1

- The main tagging rate dependences were identified w.r.t.:
  - jet E_T
  - jet N_{TRK}
  - Met *cos $\Delta\phi$(met,jet)



asymmetric distribution due to different concurring processes, HF+jets EWK+jets

allows to track sample composition changes with MET

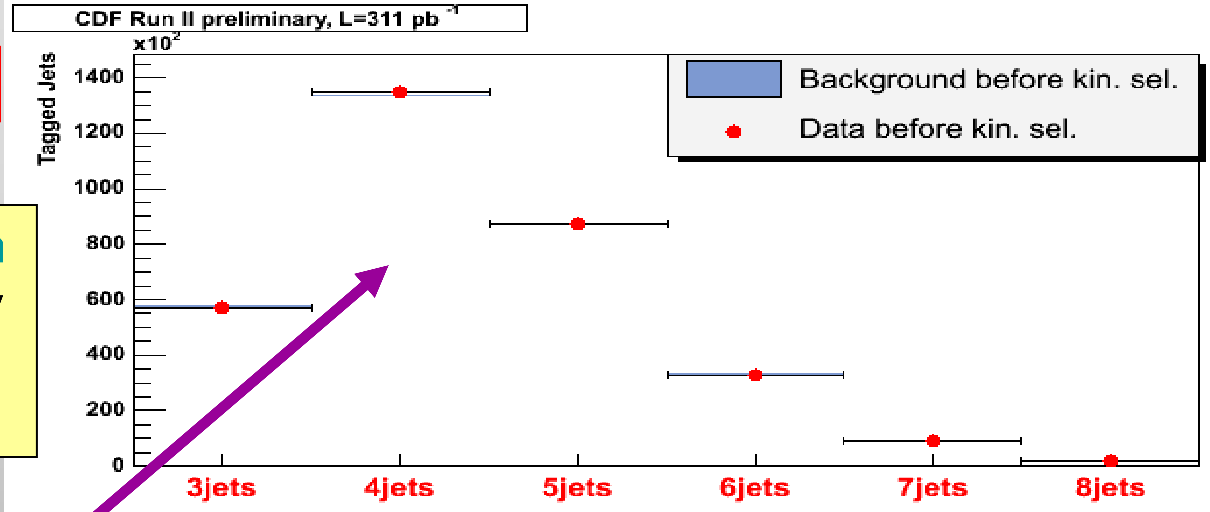
3-d (E_T, N_{TRK}, Met_{PRJ}) Positive Tagging Matrix constructed on 3 (E_T^{L5}>15GeV, |η|<2.0) jet data events. Use it to extrapolate the tag rate to higher jet multiplicities

The matrix bins are chosen to best fit the tagging dependences and to avoid low stats/undefined bins:

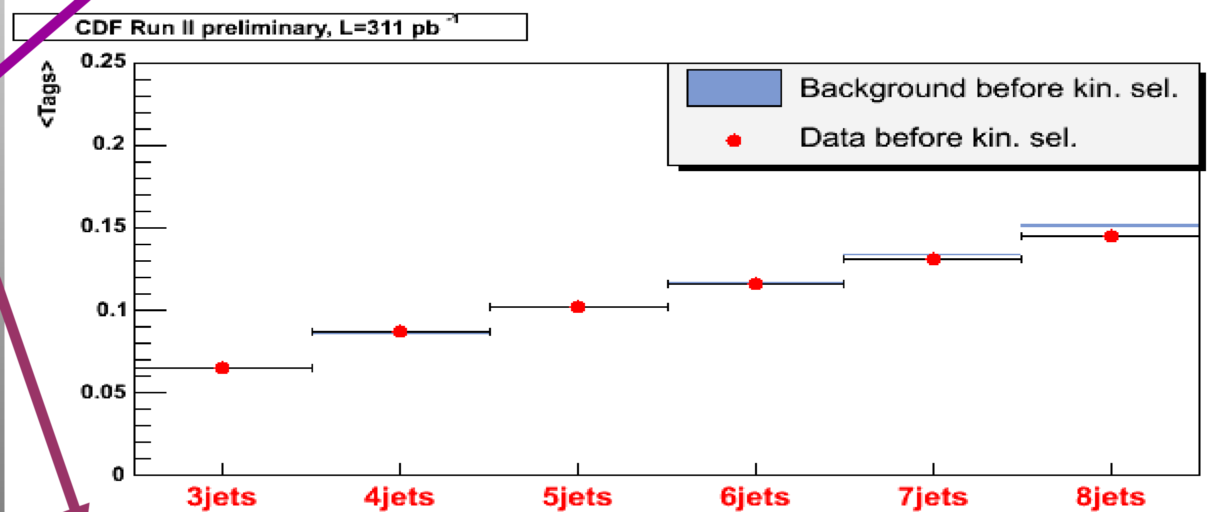
- 3 E_T bins: 15 ÷ 40, 40 ÷ 70, ≥ 70;
- 11 N_{TRK} bins: 2 ÷ 12;
- 10 Met_{PRJ} bins: ≤ -40, -40 ÷ -20, -20 ÷ -10, -10 ÷ 0, 0 ÷ 10, 10 ÷ 20, 20 ÷ 30, 30 ÷ 40, ≥ 40.

Mtx Check #1

Extrapolate the tag rate from **3 jet to higher jet multiplicity** events, before kinematical selection.



The agreement between observed and matrix-predicted positive tagged jet is good for all jet multiplicities



N jet	3	4	5	6	7	8
Nevts	879,187	1,553,525	859,543	284,062	68,628	13,237
N tgbl jets	1,781,645	4,197,888	2,877,148	1,134,089	318,410	69,966
Obs + tags	57,314	135,056	87,332	32,914	8,992	1,914
Exp + tags	$57,314 \pm 233$	$133,275 \pm 546$	$87,156 \pm 370$	$33,184 \pm 149$	$9,147 \pm 43$	$2,000 \pm 10$
					Total observed + tags:	324,008
					Total expected + tags:	$322,575 \pm 1,337$

Kinematical Selection

lepton-Id SF applied

Clean up selection:

- **Tight leptons (e/μ) veto** (no overlap w/ other L+J top analyses)
- **Trigger requirement simulation** (for MC)
- **Vertex requirements**:
  - $|Z_{vert}| < 60$ cm
  - $|Z_{vert}^{jet} - Z_{pvert}| < 5$ cm
  - $N_{vertices}(Q>12) \geq 1$

N evt	$MC_{\tau+jets}$	MC_{incl}	MJ
Good Run	1,021,924	1,021,924	4,249,644
$BR(t\bar{t} \to \tau + jets)$	149,323	–	–
Trigger	82,200	647,365	4,249,644
Nlep = 0	78,084	583,697	4,245,940
$\lvert Z_{vert}\rvert < 60\ cm$	75,035	559,342	4,051,242
$\lvert Z_{pvert} - Z_{jvert}\rvert < 5\ cm$	74,912	558,494	3,897,756
$N_{vert}(Q=12) \geq 1$	74,904	558,528	3,897,755

Optimize the kinematical selection in order to minimize the relative statistical error on xsec using both the expected amount of tags for inclusive ttbar and background (from matrix)

Optimization procedure:

- **Start by selecting ≥4 jets (matrix is computed with =3 jet events)**
- **Scan different sets of requirements (metsig, A, minΔφ)**
- **Calculate the amount of expected bkg tags for a given cut set**
- **Instead of N_{obs}^{tag} use $N^{tag}_{mc} + N^{tag}_{exp}$**

$$\sigma_{ttbar} = \frac{N_{obs}^{tag} - N_{exp}^{tag}}{\varepsilon_{kin} \cdot \varepsilon_{tag}^{ave} \cdot L}$$

Choose the set of cuts that minimizes the expected (stat. only) relative error on xsec

Kinematical Selection – cont'd

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

#	cut set			MC^{evt}_{inc}	MJ^{evt}	MC^{tag}_{inc}	BKG^{tag}	$S/\sqrt{N}$	$\frac{\sigma_{xsec}}{xsec}$
1	4.00	0.00	0.40	93	597	73	67 ± 3	8.91	17.47%
2	4.00	0.00	0.50	82	461	65	46 ± 3	9.55	17.48%
3	4.00	0.01	0.40	90	549	71	63 ± 4	8.93	17.60%
4	4.00	0.01	0.50	80	426	63	43 ± 3	9.60	17.60%
5	4.00	0.02	0.40	85	494	68	58 ± 3	8.91	17.84%
...	...	...	...	...	...	...	...	...	...

Extra acceptance come from e/μ + jets ttbar events failing the tight lepton identification requirements

before tagging

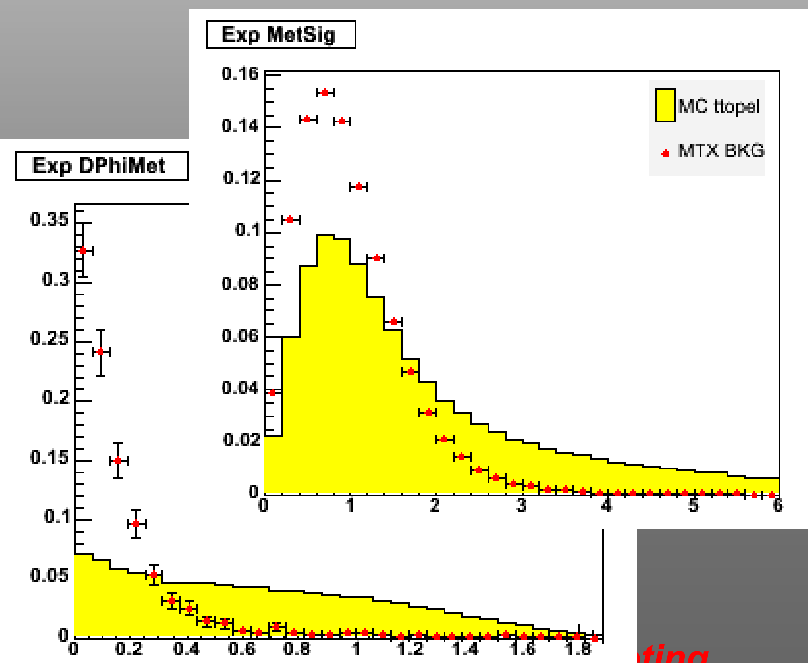
N evt	$MC_{(\tau+jets)}$	$MC_{Incl.}$	MJ
Total	149,323	1,021,924	4,249,644
Prereq	74,904	558,528	3,897,755
$N_{jet} \geq 4$	72,708	549,138	2,781,788
$\not{E}_T/\sqrt{\Sigma E_T} \geq 4$	29,830	78,145	3,996
min $\Delta\phi(\not{E}_T, jets) \geq 0.4$	19,079	49,848	597
in 311 pb^{-1}:	35.41	92.63	597

S/N $\tau + jets$: $N^{\tau}_{mc}/(N^{data}_{obs} - N^{incl}_{mc}) = 0.07$

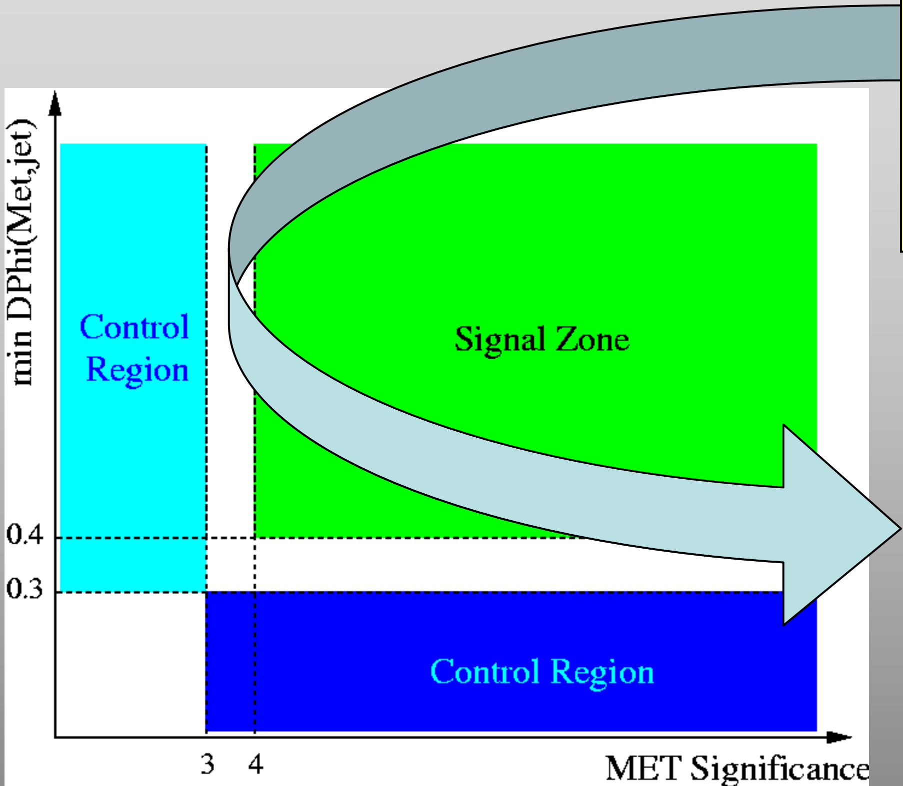
S/N Inclusive: $N^{incl}_{mc}/(N^{data}_{obs} - N^{incl}_{mc}) = 0.18$

after tagging

min $\Delta\phi(\not{E}_T, jets) \geq 0.4$	19,079	49,848	597
≥ 1 tag	11,691	30,511	106
in 311 pb^{-1}:	21.7	56.7	106

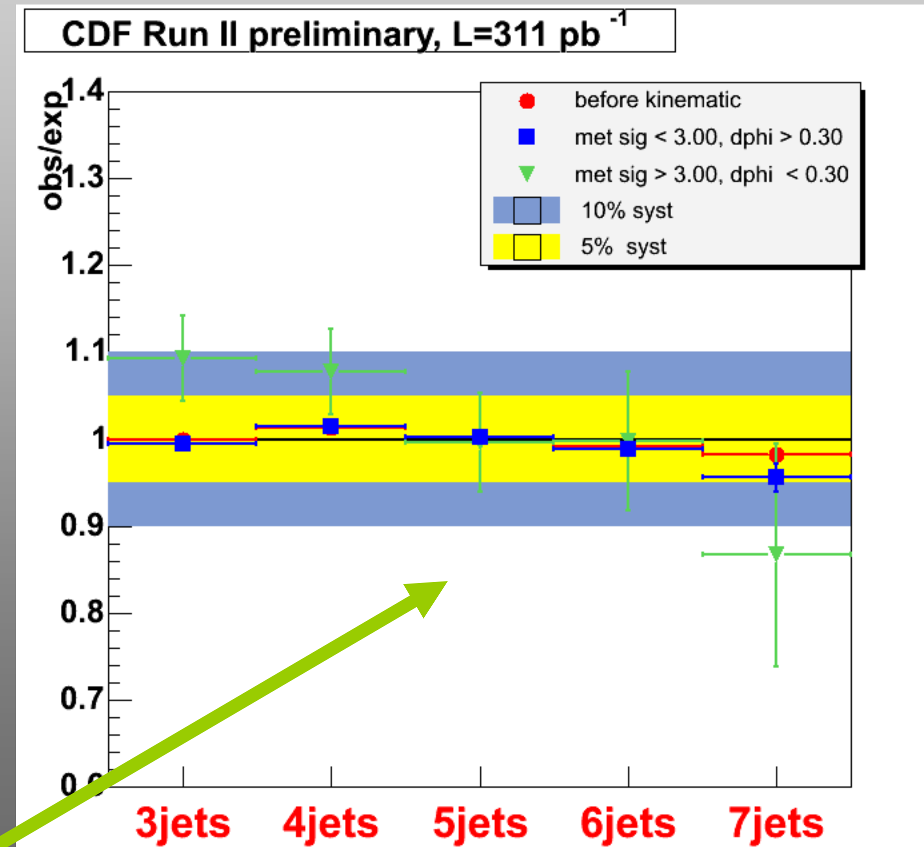


More on Matrix checks:



Once we have an optimized kin selection:

The tagging matrix background predictions can be checked in control samples obtained from multi-jet data itself:



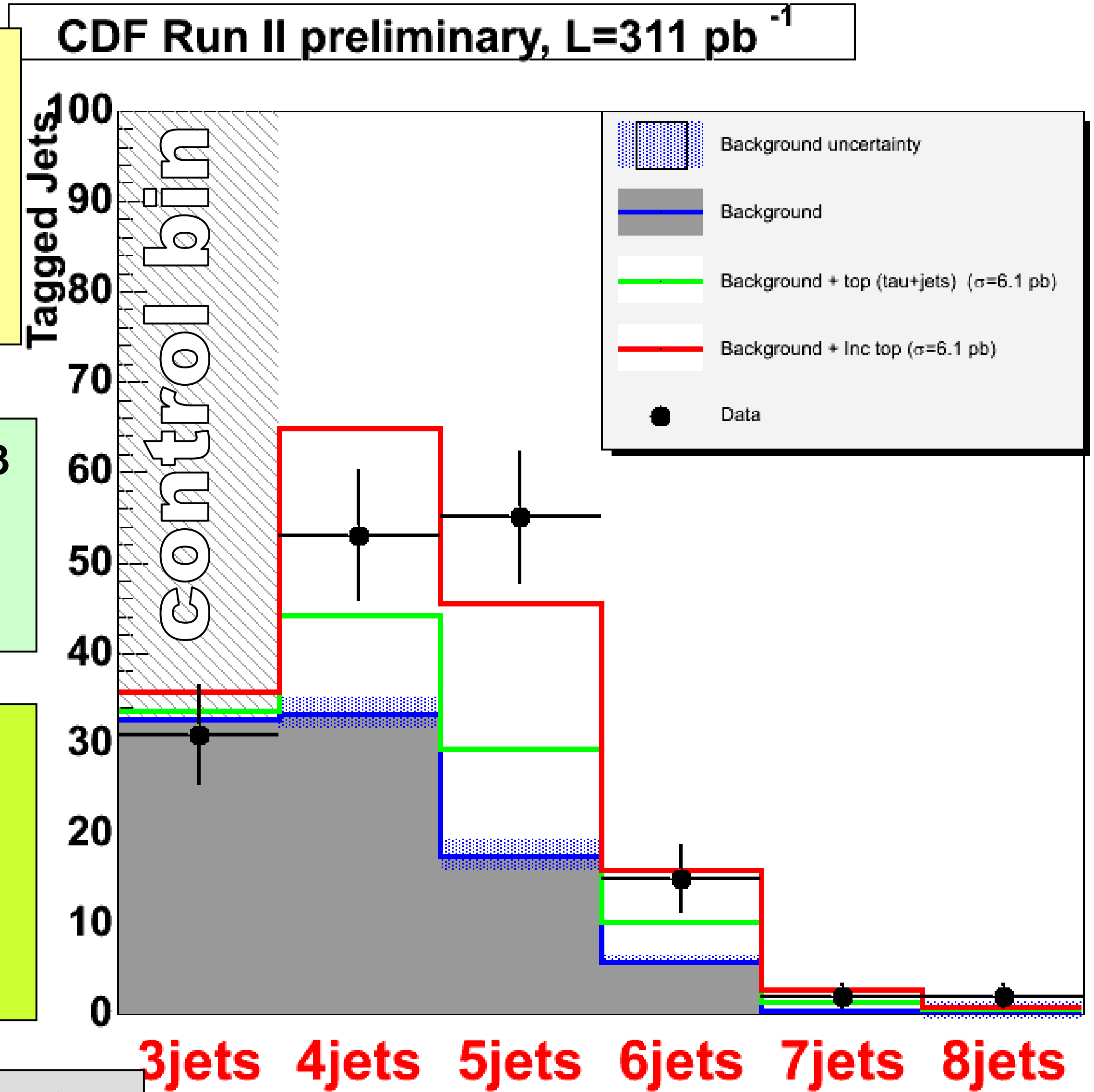
- **data before kinematical selection**
- **data w/ met sig < 3 and min$\Delta\phi$ > 0.3**
- **data w/ met sig > 3 and min$\Delta\phi$ < 0.3**

The matrix performs well in the control samples, the discrepancies in terms of the ratio obs/exp tags being well conservatively limited at 10 %.

Kin sel + ≥1 tag Sample:

Once we feel confident about our matrix parametrization we can look at its prediction in the data sample after kinematical selection and compare it with SecVtX tagged data.

- N jets(E_T^{L5} ≥15; $|\eta|<2.0$) ≥ 3
- $\not{E}_T / \sqrt{\sum E_T} \geq 4.00$
- min $\Delta\phi$(met,jet) ≥ 0.4 rad



- 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!

Kin sel + ≥ 1 tag Sample:

Let us see in deeper details which ttbar decay channel mainly contributes to the signal we expect:

We computed the Monte Carlo positive tags expectations for each decay channel as a function of the event jet multiplicity; Remember tight leptons are rejected!

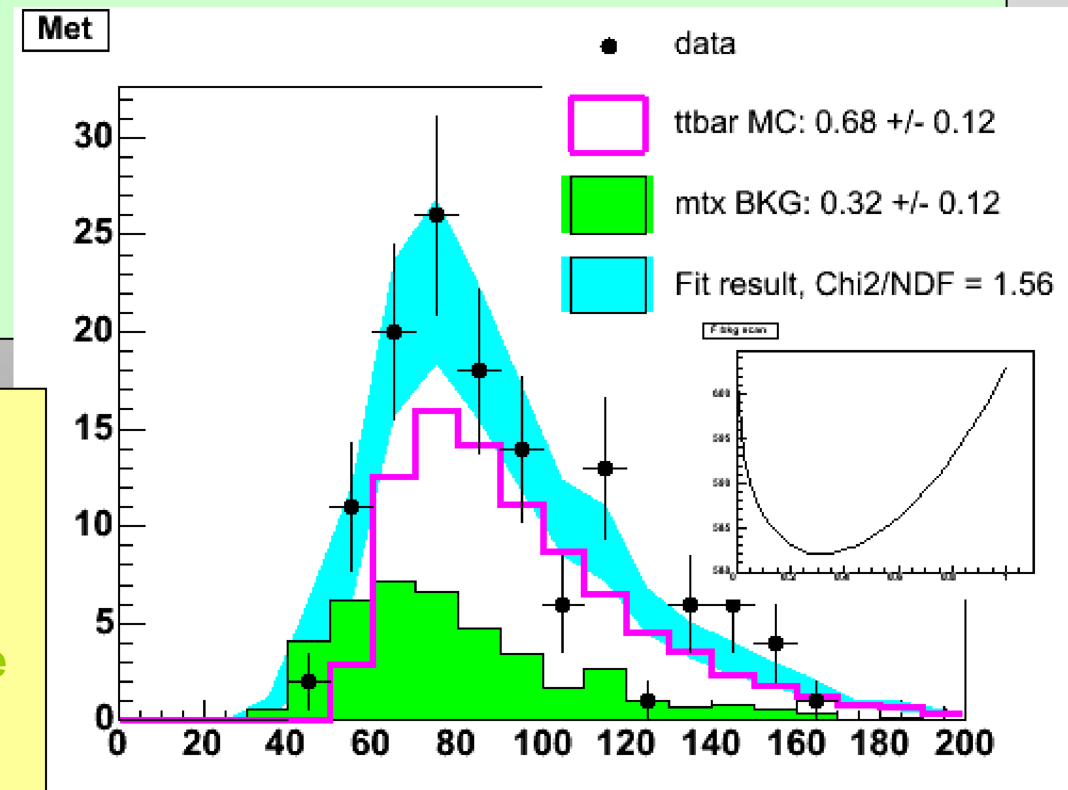
- N jets(E_T^{L5} ≥15; $|\eta|$<2.0) ≥ 3
- $\not{E}_T / \sqrt{\sum E_T} \geq 4.00$
- min $\Delta\phi$(met,jet) ≥ 0.4 rad

N jets	3	4	5	6	7	8
Monte Carlo contributions						
ee	0.08	0.41	0.18	0.04	0.02	–
$e\mu$	0.06	0.31	0.12	0.04	0.01	–
$\mu\mu$	0.02	0.06	0.01	0.01	–	–
$e\tau$	0.12	0.94	0.38	0.11	0.03	–
$\mu\tau$	0.05	0.29	0.15	0.05	0.01	–
$\tau\tau$	0.07	0.57	0.25	0.04	0.01	–
$e + jets$	0.66	6.64	8.68	3.34	0.74	0.20
$\mu + jets$	1.06	11.85	6.55	2.03	0.37	0.07
$\tau + jets$	0.99	11.05	11.76	4.36	0.97	0.18
$all - had$	0.01	0.08	0.13	0.15	0.04	0.03
$inclusive$	3.10	32.21	28.23	10.20	2.22	0.48

Checking F_{top} using 2-c fits

We can cross-check the excess we attribute to ttbar production by looking to kinematical variables.

In particular we can use the positive tagging matrix to extract background shapes.



Then fit data distribution after kin sel + ≥1 tag to the sum of:

- **Inclusive ttbar template**
- **Matrix extracted bkg template**

And extract the relative fractions of signal and background in the data

Used binned likelihood fits:

$$\begin{cases} L = -2 \cdot \sum_{i=1}^{Nbins} N_i \cdot \log(F) - F \\ where : F = f_b B_i + (1 - f_b) S_i \end{cases}$$

and checked the fitting procedure by pseudoexps.

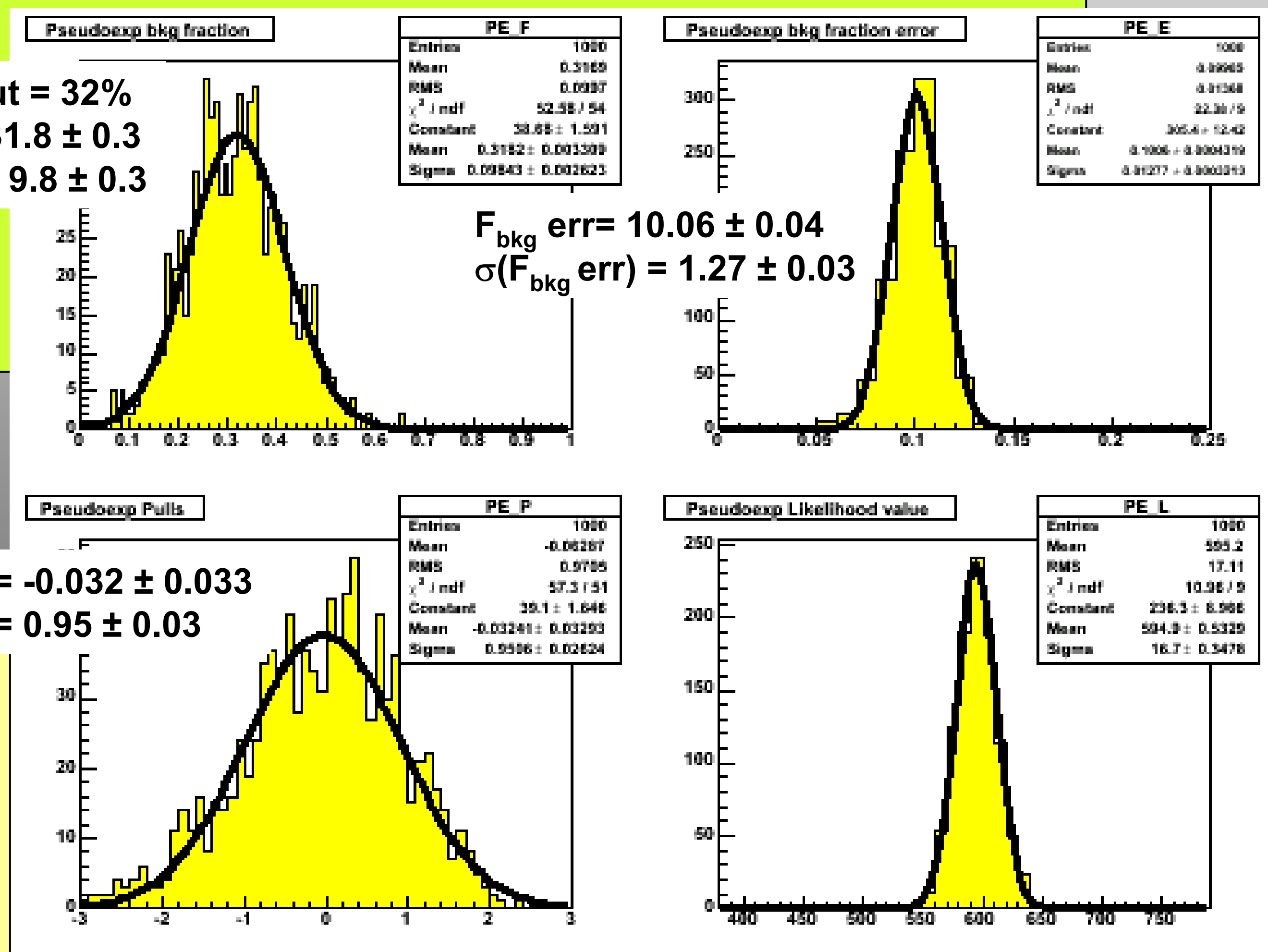
Checking the fit procedure..

Pseudo experiments:

Use the fitted background fraction to generate pseudo-experiments w/ same stats as data from the original shapes for signal and background.

Fit them with the same fitting procedure…

…and iterate 1K times



F_{bkg} input = 32%
F_{bkg} = 31.8 ± 0.3
$\sigma(F_{bkg})$ = 9.8 ± 0.3

F_{bkg} err= 10.06 ± 0.04
$\sigma(F_{bkg}$ err) = 1.27 ± 0.03

Pull mean = -0.032 ± 0.033
Pull sigma= 0.95 ± 0.03

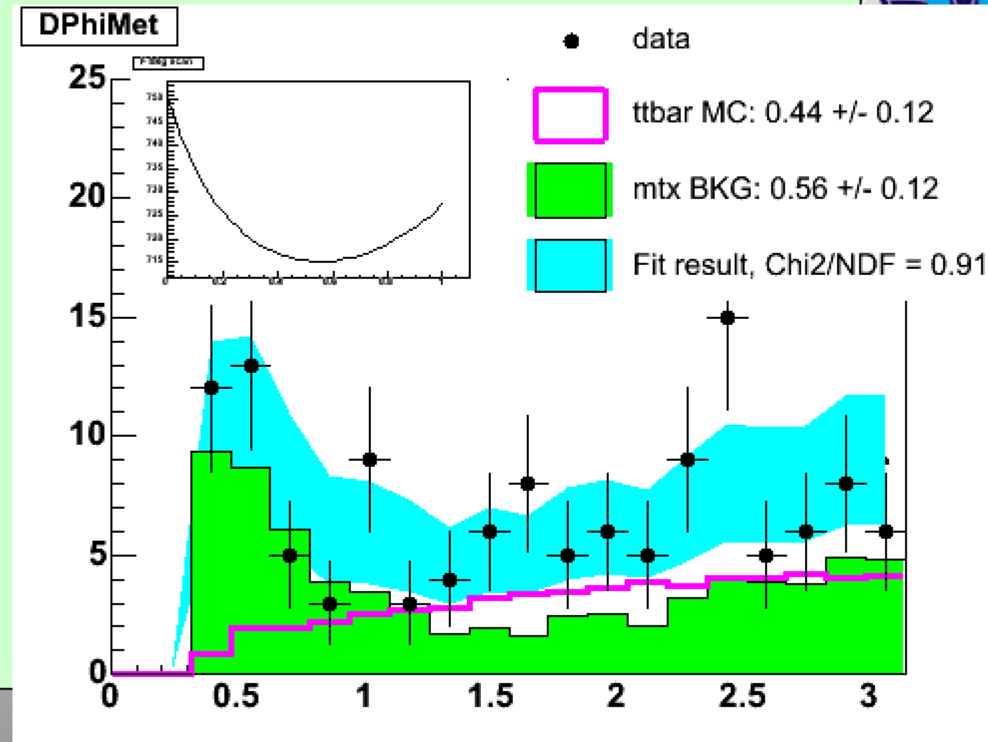
Fig refers to the MET fits

Fit technique does not show significant bias

We then fitted several other event and/or jet variables w/ the same technique.

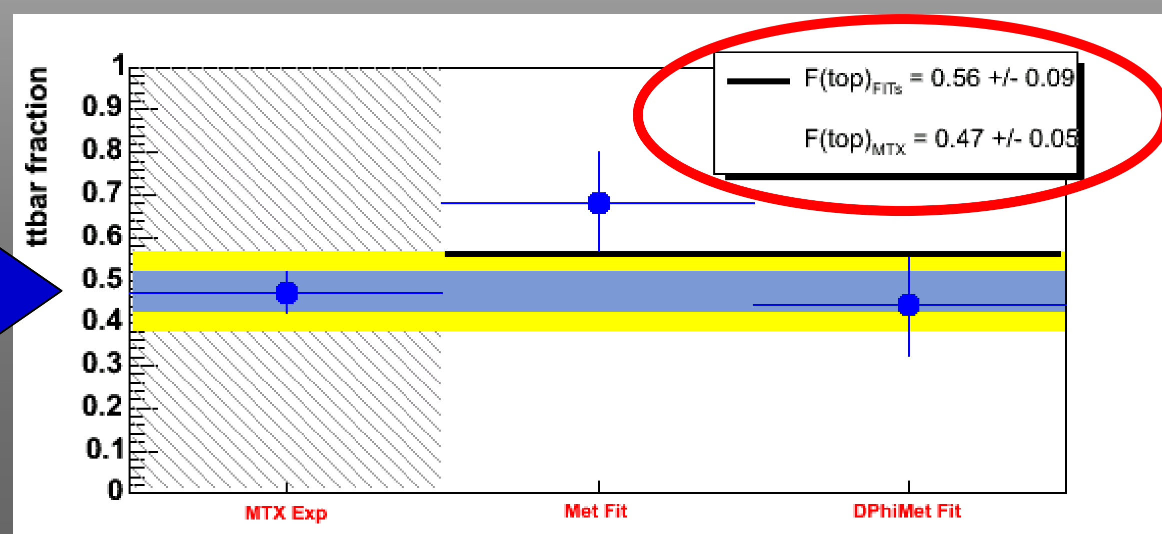
Anyway, most of them are correlated with each other.
For this reason we chose one event-variable and one jet-variable for our estimation.

We chose among the vars the ones on which the PE return on average the lowest error on the fitted fraction.



From data fits:

we found a ttbar fraction consistent with that calculated by the counting method.

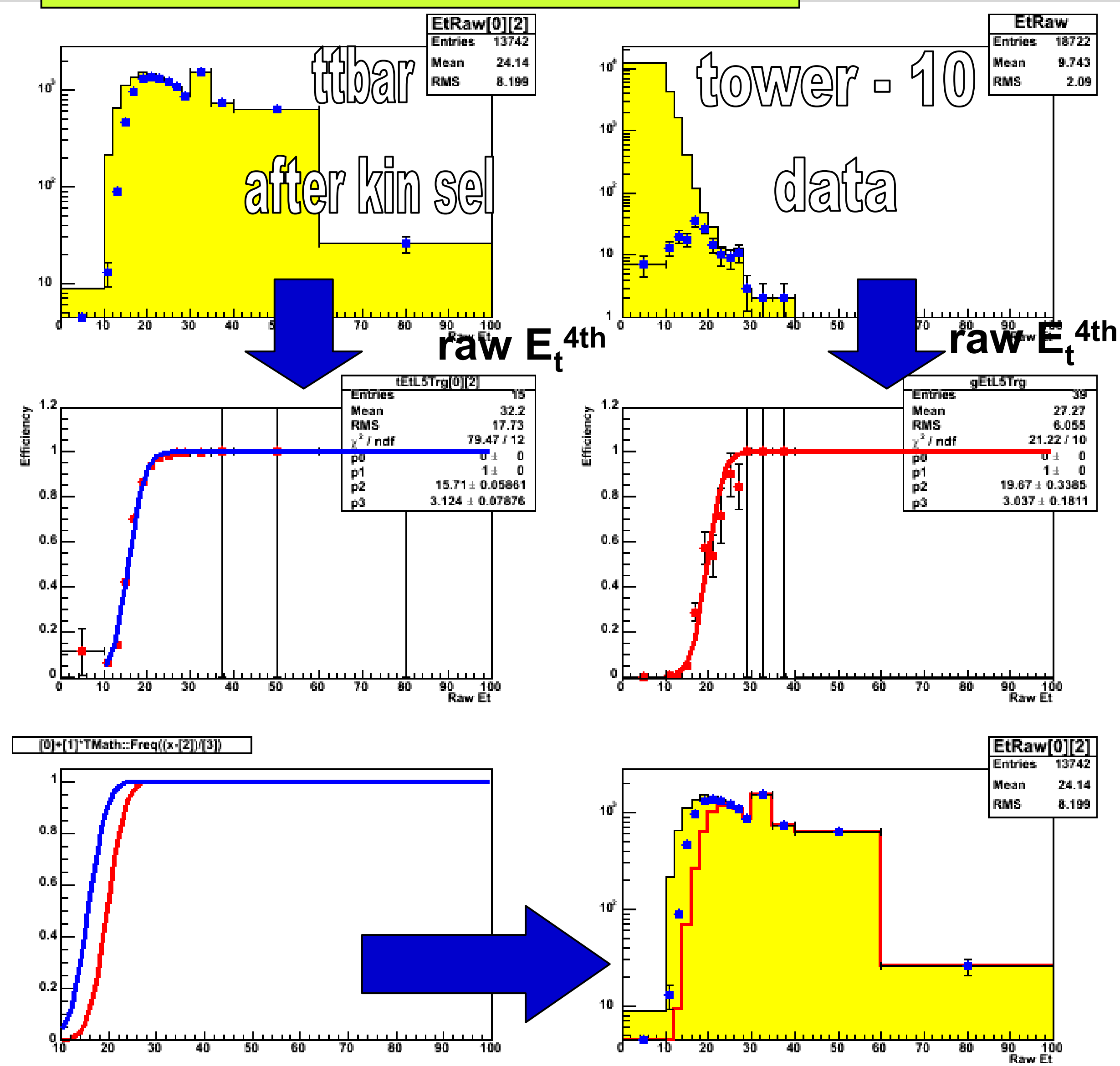


- Systematic Uncertainties -

Trigger Systematic effect:

CDF-7473

Trigger efficiency on signal events is determined using TRIGSIM++.



Need to evaluate related systematic:

comparing trigger turn-on curve (as a function of some offline variable) as returned by the simulation and as measured from Tower 10 data (same L1 as TOP_MULTI_JET).

The mismatch between turn-on curve allows to quantify the systematic effect at 17.8%

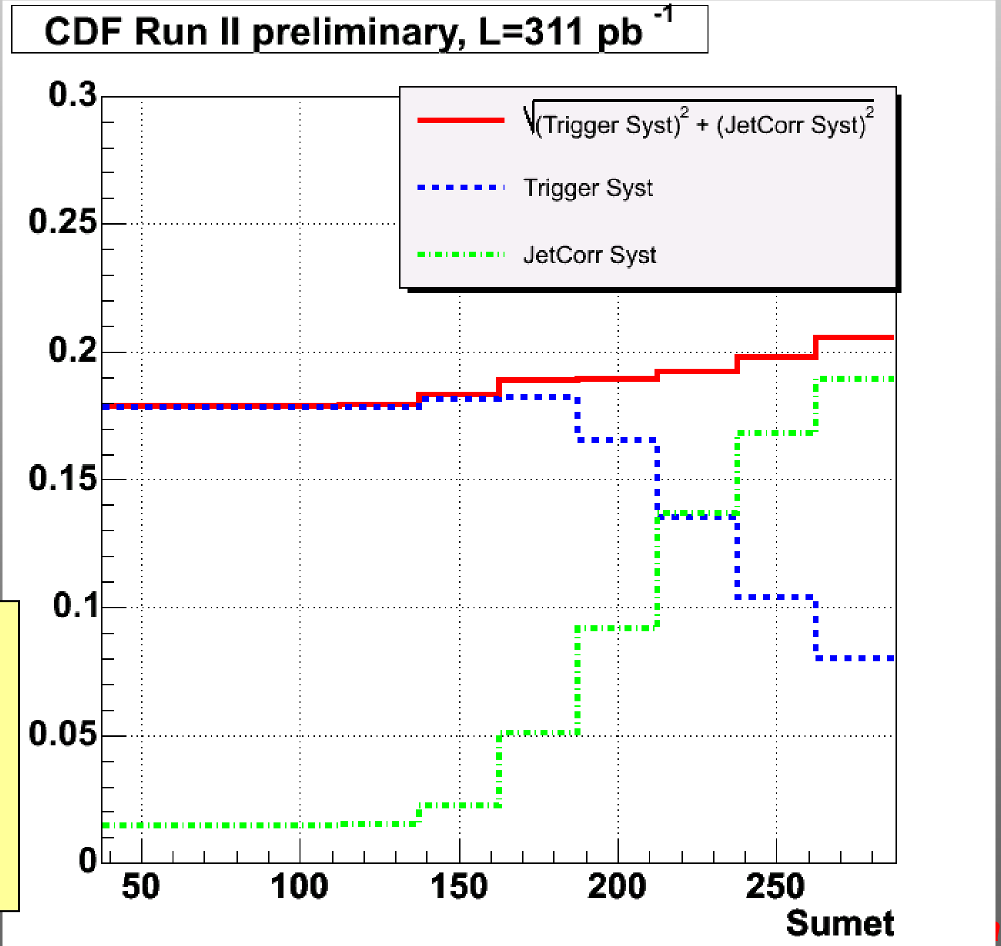
Note: the 4th offline jet is matched with the 4th L2 cluster within R=0.4 in order to preserve energy hierarchy.

Trigger Systematics – cont'd

We studied the trigger systematic effect as a function of an extra kinematical cut on ΣE_T in the range 50-275 GeV in order to possibly reduce it

The trigger systematic is indeed found to decrease with an extra cut on ΣE_T around 200-225 GeV but on the other hand the effect is compensated by the increase of the systematic effect related to jet energy corrections.

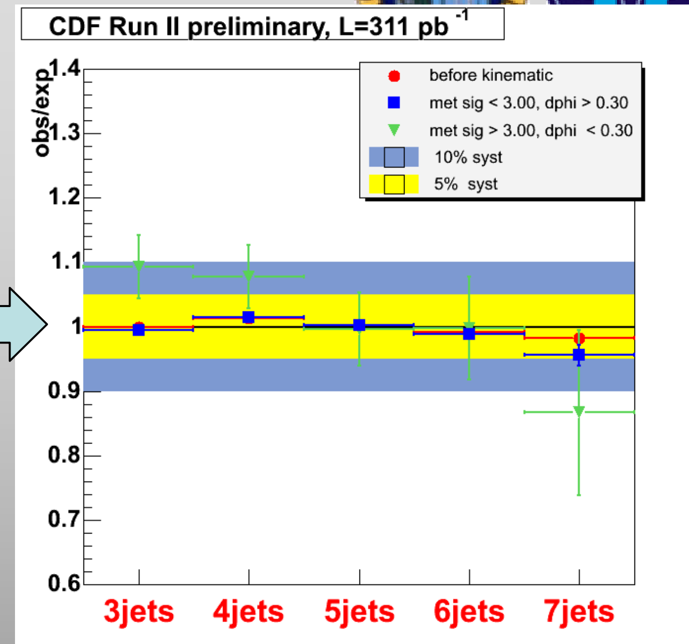
ΣE_T cut	Trig Syst [%]	JetCorr Syst [%]	Total [%]
50	17.82	1.48	17.88
75	17.82	1.48	17.88
100	17.82	1.48	17.88
125	17.84	1.54	17.91
150	18.16	2.23	18.30
175	18.20	5.10	18.90
200	16.55	9.19	18.93
225	13.50	13.68	19.22
250	10.41	16.81	19.77
275	7.98	18.93	20.54



The sum in quadrature of trigger and jet energy correction systematics is found to increase w/ the applied ΣE_T cut.

Systematic Uncertainties:

- **Background prediction:** from obs/exp comparison in control samples we quote **10%.**



- **Luminosity:** 4.4% from CLC acceptance and 4% from inelastic xsec: **6%**

- **PYTHIA/HERWIG** generator dependence: **8.2%**

$\frac{\epsilon_{HERWIG}-\epsilon_{PYTHIA}}{\epsilon_{PYTHIA}}$	PYTHIA vs HERWIG $MC_{(\tau+jets)}$	$MC_{Incl.}$
Good Run	–	–
$BR(t\bar{t} \to \tau + jets)$	1.1 %	–
Trigger	6.4 %	2.9 %
Nlep = 0	0.2 %	-0.7 %
$\lvert Z_{vert}\rvert < 60\ cm$	0.1 %	0.0 %
$\lvert Z_{pvert} - Z_{jvert}\rvert < 5\ cm$	0.0 %	0.0 %
$N_{vert}(Q=12) \geq 1$	0.0 %	0.0 %
$N_{jet} \geq 4$	0.2 %	0.1 %
$\not{E}_T/\sqrt{\Sigma E_T} \geq 4$	-0.1 %	5.8 %
min $\Delta\phi(\not{E}_T, jets) \geq 0.4$	0.0 %	0.1 %
total	7.9 %	8.2 %

- **ISR/FSR** systematic: **2 %**
- **PDF**-systematic: **1.6 %**

Systematic Uncertainties:

- **JES systematic: 1.5 %.**

It is low due to trigger effects ($N_{jets} \geq 4$ cut) **For $N_{jets} \geq 5$, JES-systs is 7.6 %.**
and due to the fact we used met/$\sqrt{\Sigma E_T}$ ratio to select our events.

- **B-tagging Scale Factor dependence: 5.5 %**

N evt MC_{incl}	standard SF	$+1\sigma$ SecVtx SF	-1σ SecVtX SF
Total	1,021,924	1,021,924	1,021,924
After Kin Sel	49,848	49,848	49,848
$\geq$ 1 tag	30,511	31,578	29,106
Tot tags	39,422	41,406	37,051
ϵ_{tag}^{evt}%	79.08 ± 0.18	83.06 ± 0.17	74.33 ± 0.20

Systematic Uncertainties:

We dropped Top Mass dependence as a syst.

Systematics uncertainty sources

Source	Method	Uncertainty
ϵ_{kin} systematics		
Trigger simulation	turn-on curves	17.8 %
Generator dependence	$\frac{\lvert\epsilon_{PYTHIA}-\epsilon_{HERWIG}\rvert}{\epsilon_{PYTHIA}}$	8.2 %
PDFs	MC reweighting	1.6 %
ISR/FSR	samples comparison	2.0 %
Jet Corrections	$\frac{\lvert\epsilon_{jetcorr,+1\sigma}-\epsilon_{jetcorr,-1\sigma}\rvert}{2\epsilon_{kin}}$	1.5 %
ϵ_{tag} systematics		
SecVtX scale factor	$\frac{\lvert\epsilon_{tag,+1\sigma}-\epsilon_{tag,-1\sigma}\rvert}{2\epsilon_{tag}}$	5.5 %
Tagging matrix systematics		
Data control samples	N_{obs}/N_{exp}	10.0 %
Luminosity systematics		
Luminosity measurement	–	6.0 %

- **Total Systematic uncertainty: 24 %**

Pre-tag iterative top subtraction

- 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 ±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:

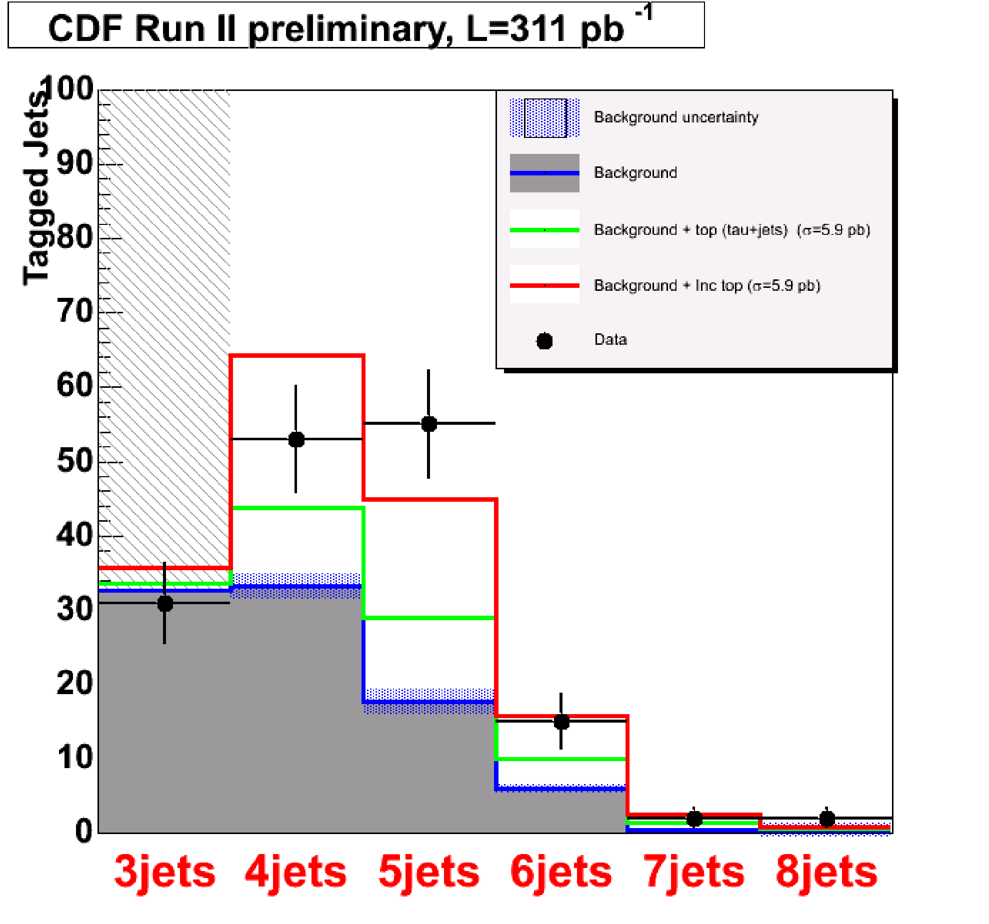
$$N'_{\exp} = N^{fix}_{\exp} \frac{N_{evt} - N^{ttbar}_{evt}}{N_{evt}} = N^{fix}_{\exp} \frac{N_{evt} - \frac{N_{obs} - N_{\exp}}{\varepsilon^{ave}_{tag}}}{N_{evt}}$$

- 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 ± 8.1 is the corrected background amount to be used for a cross section measurement.

Cross Section:

- The cross section is measured by means of a likelihood function maximization:

$$\mathcal{L} = e^{-\frac{(L-\bar{L})^2}{2\sigma_L^2}} \cdot e^{-\frac{(\epsilon_{kin}-\bar{\epsilon}_{kin})^2}{2\sigma_{\epsilon_{kin}}^2}} \cdot e^{-\frac{(\epsilon_{tag}-\bar{\epsilon}_{tag})^2}{2\sigma_{\epsilon_{tag}}^2}} \cdot e^{-\frac{(N'_{exp}-\bar{N}'_{exp})^2}{2\sigma_{N'_{exp}}^2}} \cdot \cdot \frac{(\sigma_{t\bar{t}} \cdot \epsilon_{kin} \cdot \epsilon_{tag} \cdot L + N'_{exp})^{N_{obs}}}{N_{obs}!} \cdot e^{-(\sigma_{t\bar{t}} \cdot \epsilon_{kin} \cdot \epsilon_{tag} \cdot L + N'_{exp})}$$



Variable	Simbol	Input Value	Output Value
Integrated Luminosity (pb^{-1})	$\mathcal{L}$	311 ± 18	311 ± 18
Observed Tags	N_{obs}	127	–
Expected Tags	N'_{exp}	57.4 ± 8.1	57.4 ± 8.1
Kin efficiency (%)	ϵ_{kin}	4.878 ± 0.967	4.821 ± 0.997
Ave Tag efficiency (%)	ϵ_{tag}	79.08 ± 4.34	78.94 ± 4.34

- $\sigma_{ttbar} = 5.9 \pm 1.1\ (stat)\ ^{+1.6}_{-1.1}\ (syst)$ pb

  $= 5.9^{+1.9}_{-1.6}$ pb.

Conclusions:

- We wanted to isolate the ttbar tau+jets signal from multijet triggered data in the sample after kin sel + ≥ 1 SecVtX tag.
- We set up a Method-I analysis constructing a positive tagging matrix able to predict the amount of background tags in a given data sample with an uncertainty of 10%
- We optimized the kinematical selection using the matrix predicted background information by minimizing the expected statistical uncertainty in a xsec measurement. We ended up with a selection showing extra acceptance mainly from "dirty" e/μ+jets events.
- By using tag counting and kinematical distribution fits, 50% of the final sample was attributed to inclusive ttbar production
- The total systematic uncertainty estimated to be 24%, and was found to be mainly driven by trigger simulation systematics.
- With all these ingredients we measured a cross section of:

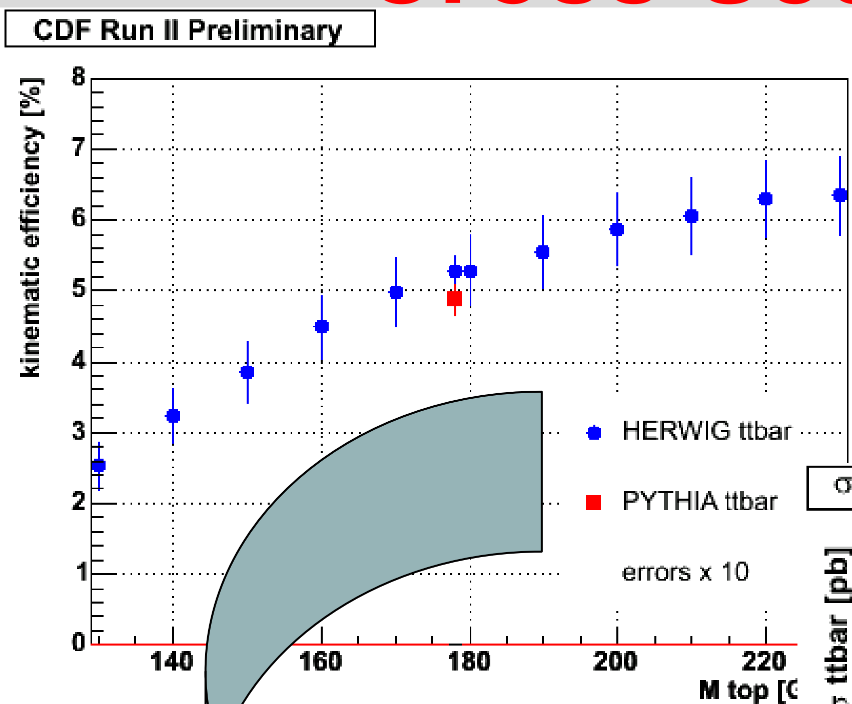
- $\sigma_{ttbar} = 5.9 \pm 1.1\ (stat)\ ^{+1.6}_{-1.1}\ (syst)\ pb$

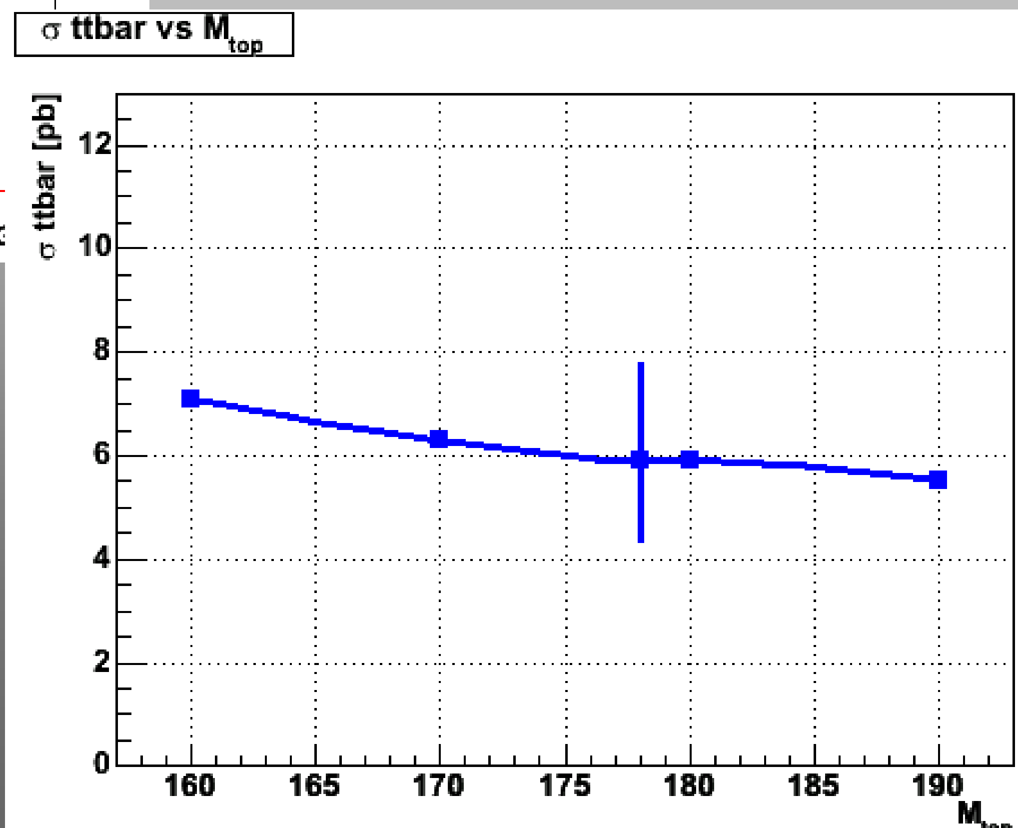
  $= 5.9\ ^{+1.9}_{-1.6}\ pb.$

- The plan is to go for the complete blessing in two weeks.

Backup Slides

Cross Section vs M_{top}

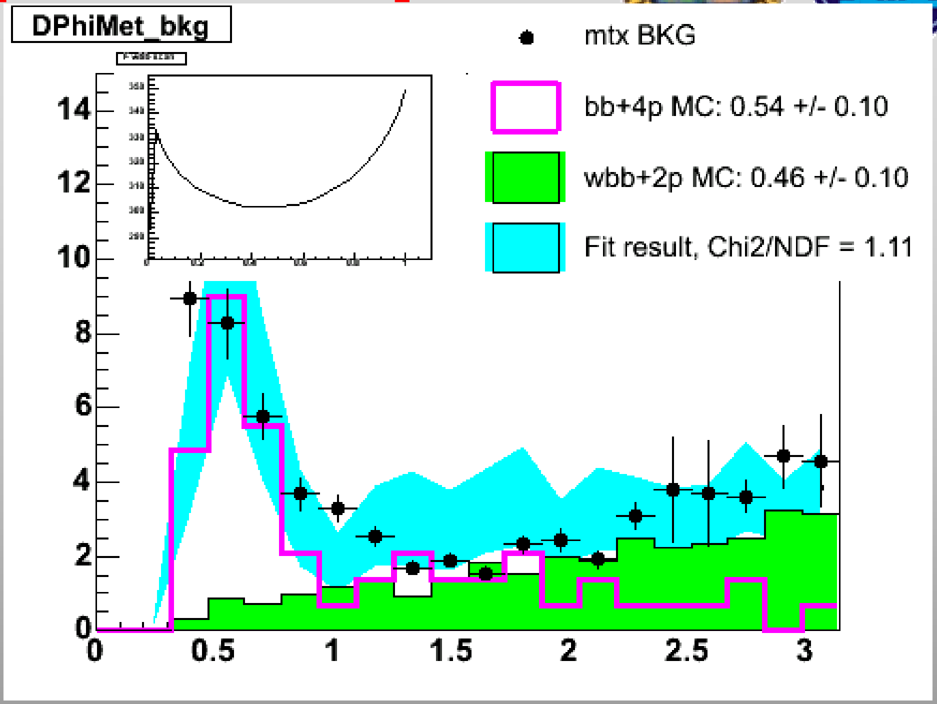




Hints on bkg sample composition

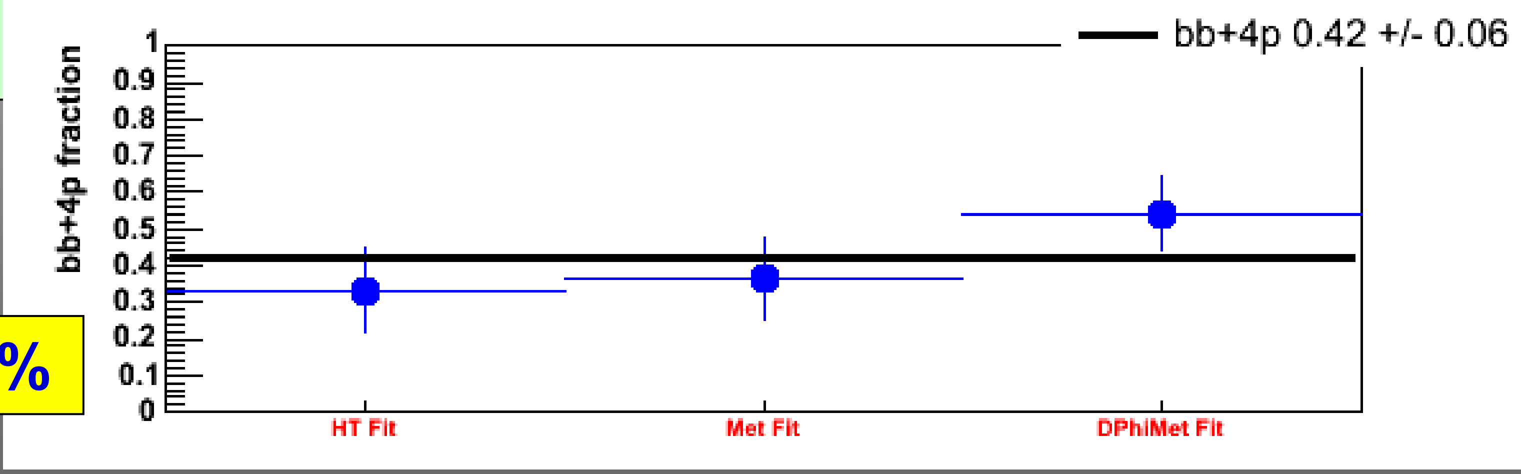
We can do more: we have the background shapes extracted from the tagging matrix information, we can fit them to the sum of two Alpgen Monte Carlo templates for the processes we expect to populate our signal region.

- **Wbb+2P**
- **bb+4P**



From bkg fits:

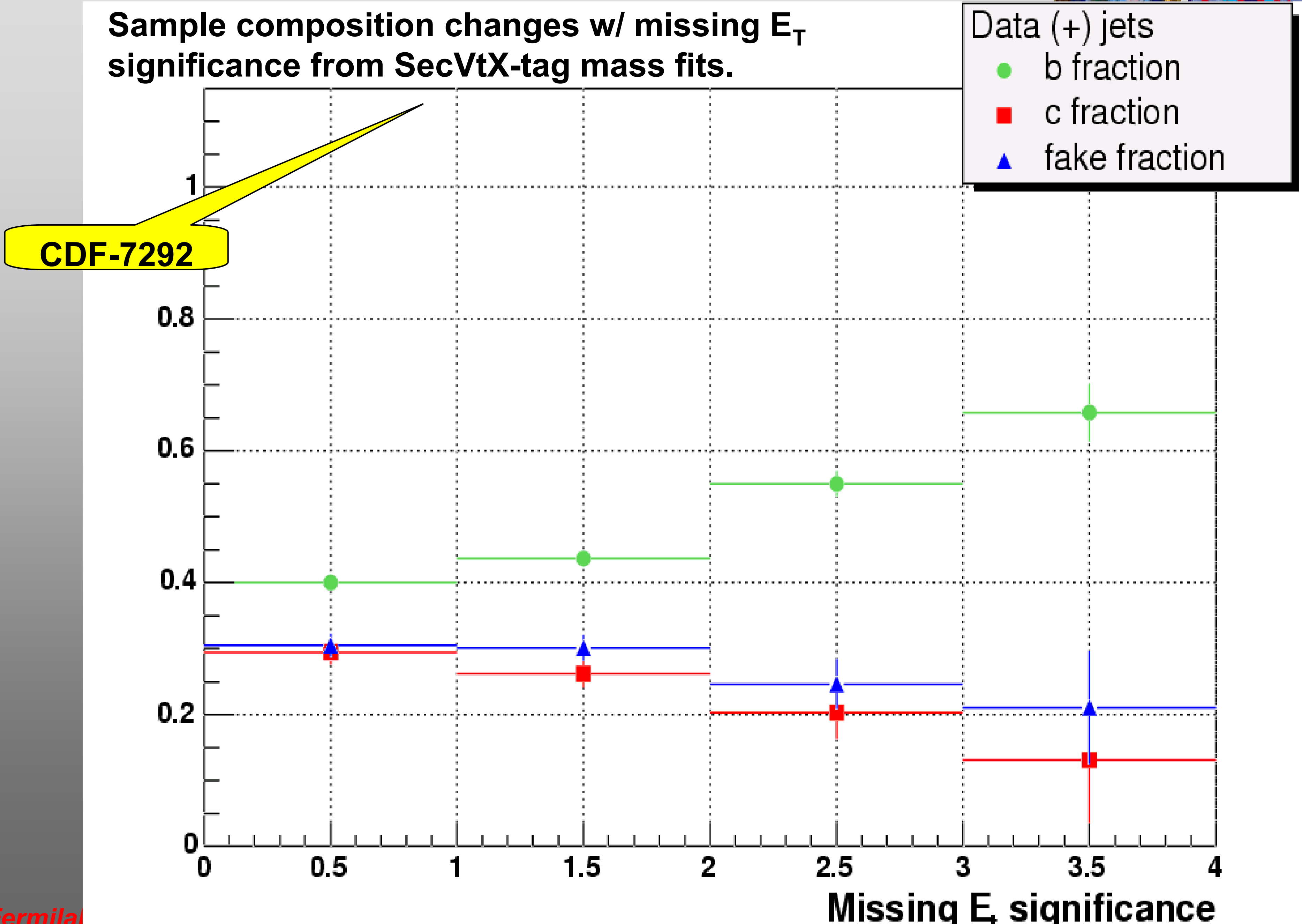
we found a bb fraction:



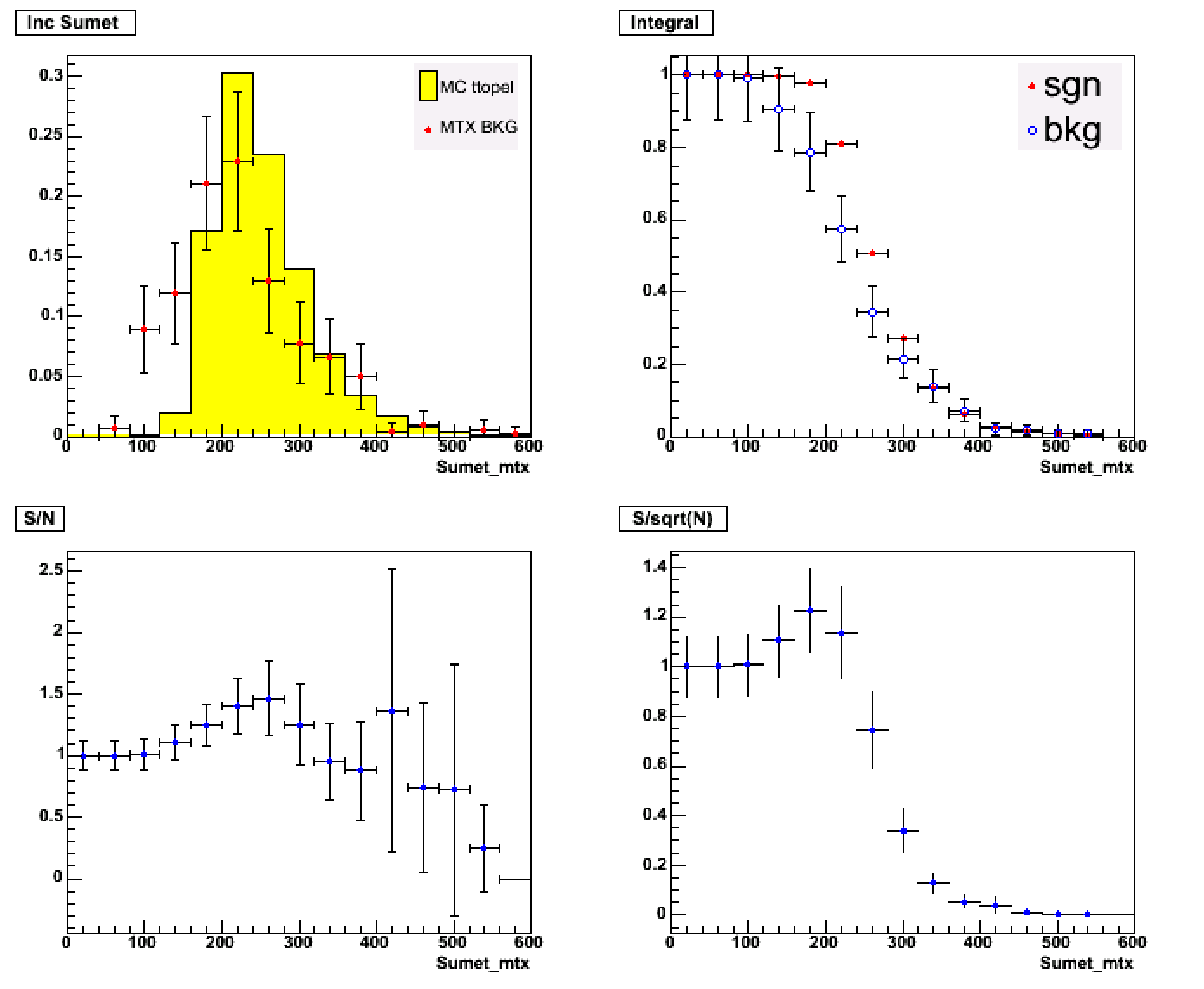
F_{bbar} ~ 42%

After kin sel + ≥1 tag data: 50% top + ~21% bb + ~ 29% Wbb

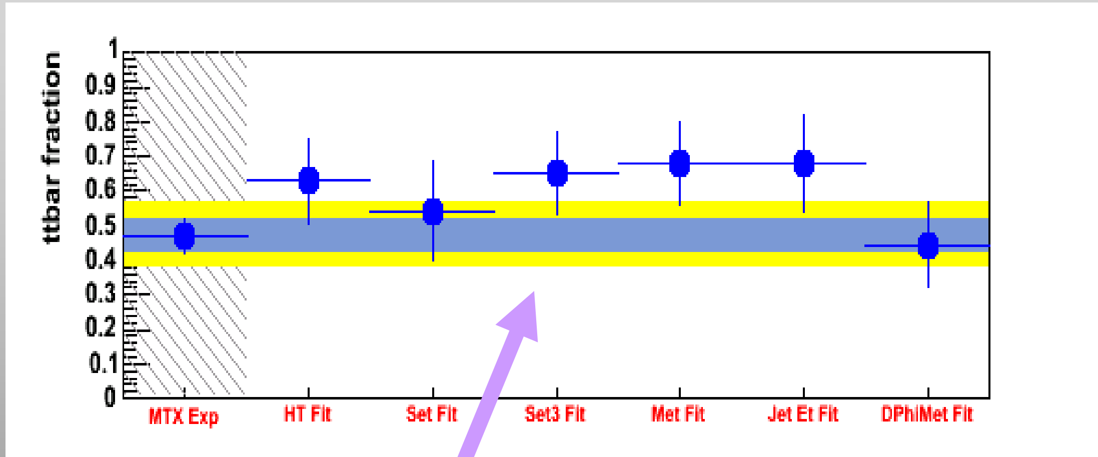
Sample composition changes w/ missing E_T significance from SecVtX-tag mass fits.

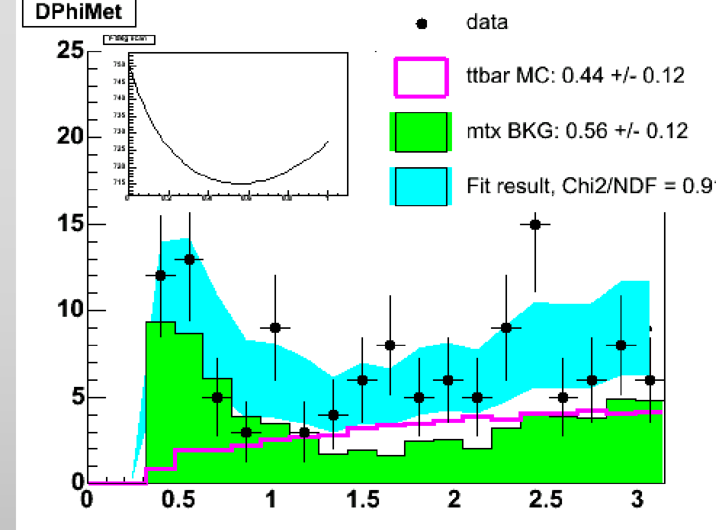


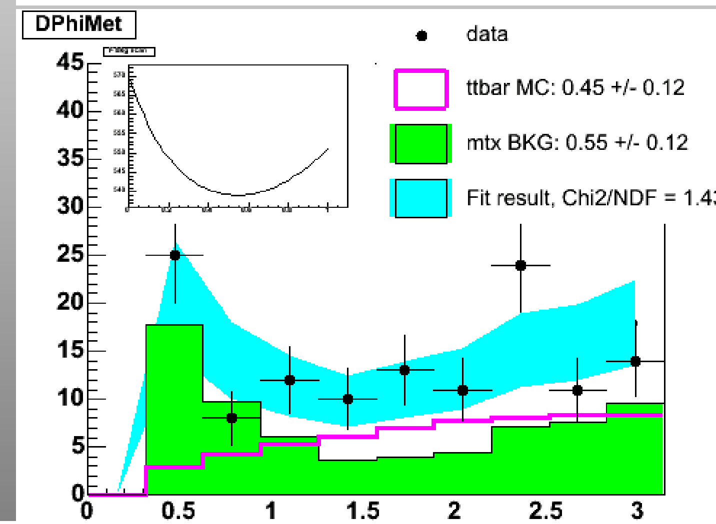
ΣE_T cut optimization after having applied the kinematical selection and the additional requirement for at least one positive tagged jet.



More details on Kinematical Fits:

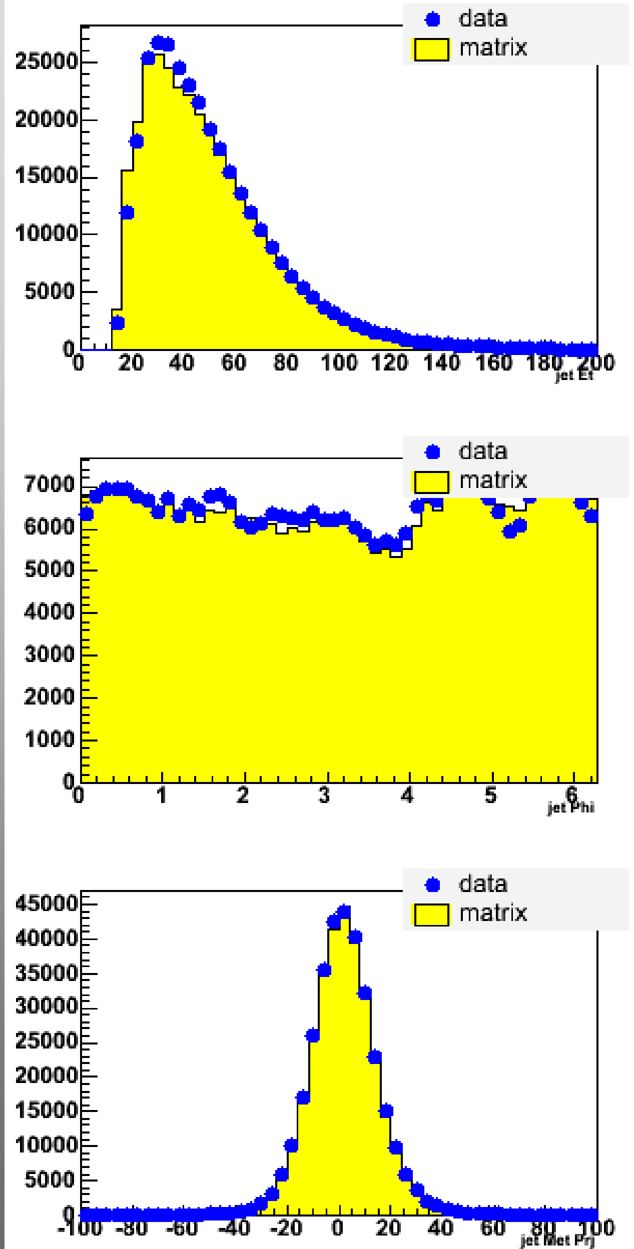






All fits results and PE output

Fitted variable	F_{bkg}	$F_{t\bar{t}}$	$\sigma(F_{bkg})^{PE}$	$pull^{PE}$ mean	$pull^{PE}$ sigma
H_T	0.372 ± 0.122	0.628 ± 0.122	0.103 ± 0.012	-0.09 ± 0.03	0.93 ± 0.03
ΣE_T	0.457 ± 0.143	0.542 ± 0.143	0.119 ± 0.012	-0.03 ± 0.03	0.94 ± 0.03
ΣE_T^3	0.349 ± 0.118	0.651 ± 0.118	0.108 ± 0.011	-0.11 ± 0.03	0.97 ± 0.03
$\not{E}_T$	0.317 ± 0.120	0.682 ± 0.120	0.101 ± 0.012	-0.03 ± 0.03	0.95 ± 0.03
jet E_T	0.317 ± 0.140	0.682 ± 0.140	0.127 ± 0.008	0.03 ± 0.03	0.86 ± 0.03
jet min $\Delta\phi(\not{E}_T, jet)$	0.555 ± 0.121	0.444 ± 0.121	0.126 ± 0.008	0.05 ± 0.03	0.95 ± 0.03

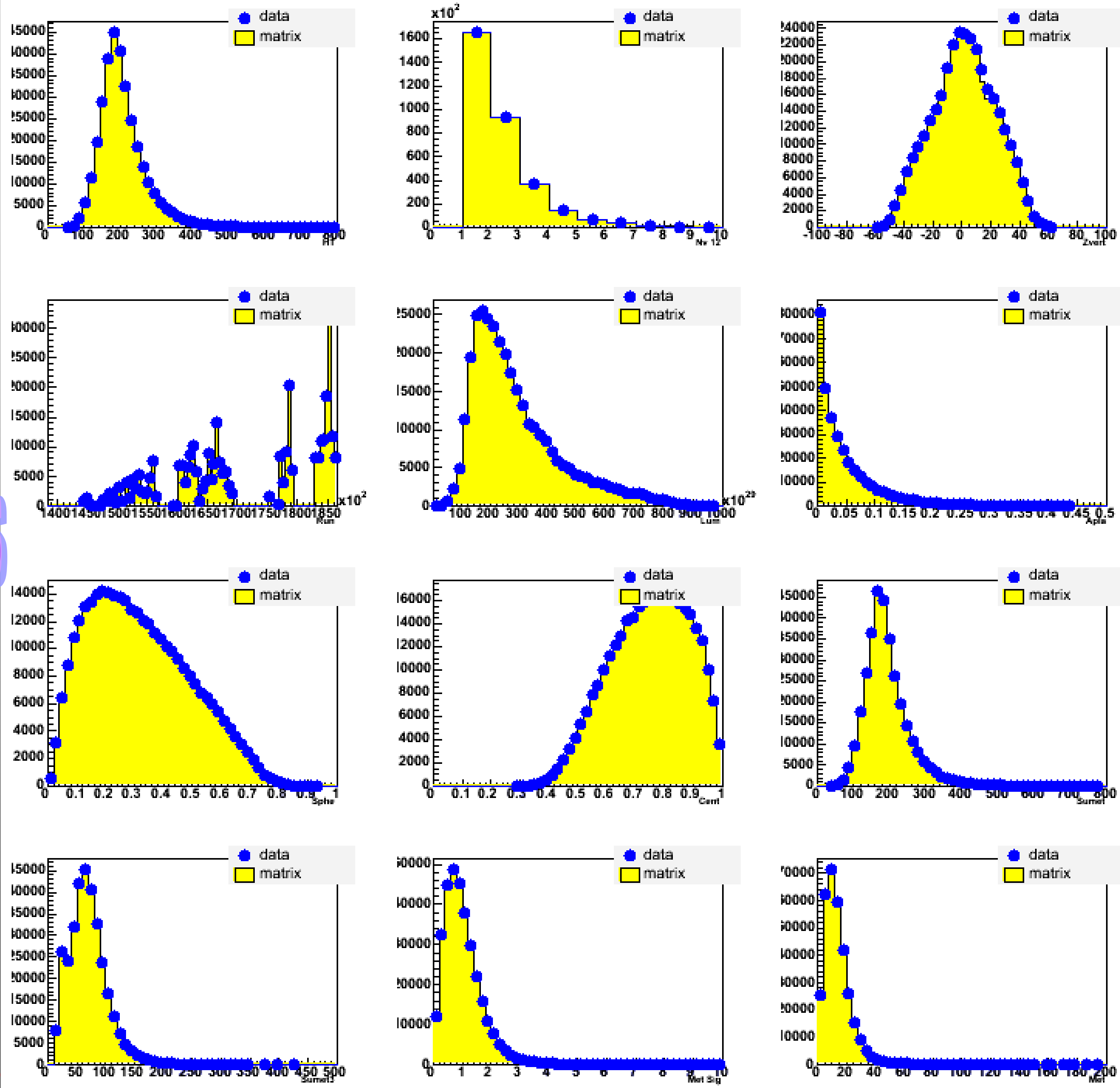


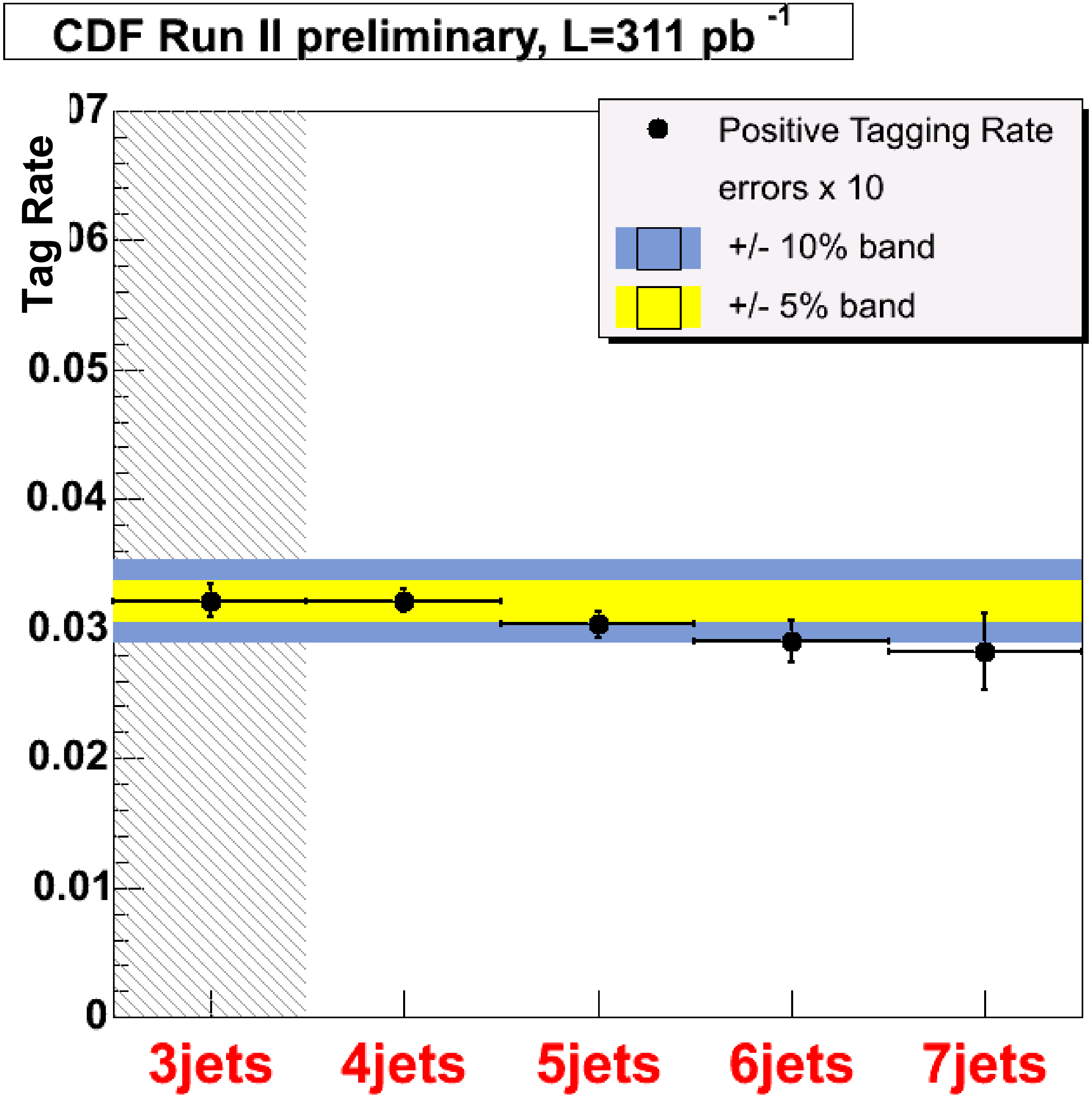
Tagging matrix based kinematical distribution compared to data ones before kinematical selection

Jet variables

Tagging matrix based kinematical distribution compared to data ones before kinematical selection

Evt variables





Positive tagging rate as a function of jet multiplicity in the data sample before kinematical selection.