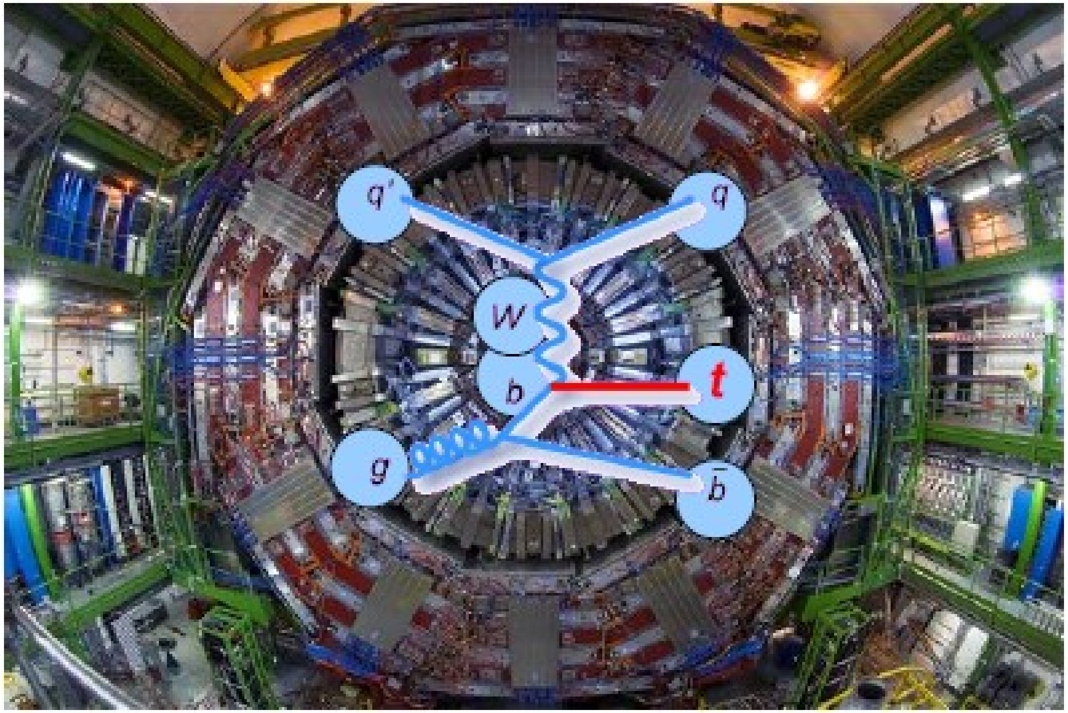
Single top in CMS: an early data strategy

A.Giammanco (Louvain University & FNRS)



Mostly based on **CMS-PAS-TOP-09-005**,

"Prospects for a measurement of the single top t-channel cross section in the muon channel, with the first 200 pb^{-1} of CMS data at 10 TeV"

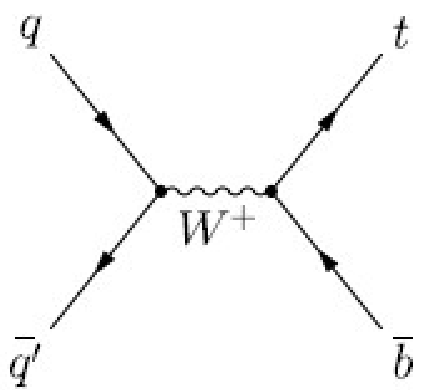
Outline

- Intro
  - Signal definition
  - Single top and CKM
  - From Tevatron to LHC
- The analysis
  - Event selection
  - Estimating QCD *in situ*
  - Signal extraction
  - Systematics
- Conclusions

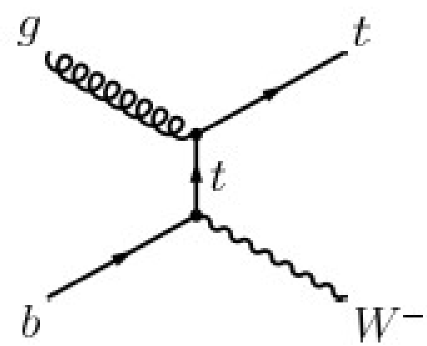
Acknowledgements:
Julia Bauer, Jeannine Wagner-Kuhr, Dmitri Konstantinov, Mojtaba Mohammadi, Fabio Maltoni, Rikkert Frederix

Main references of this talk:
- CMS-PAS-TOP-09-005
- (internal) CMS Analysis Note 2009/069
- (internal) CMS Analysis Note 2009/024
- *"Is V_{tb} ~ 1?"*, Eur.Phys.J. C49 (2007) 791

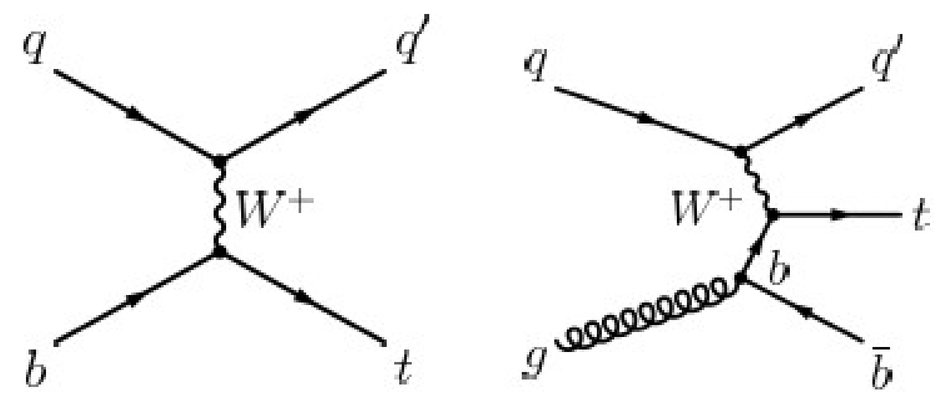
What we talk about, when we talk about single top



s channel (5 pb)



tW (29 pb)

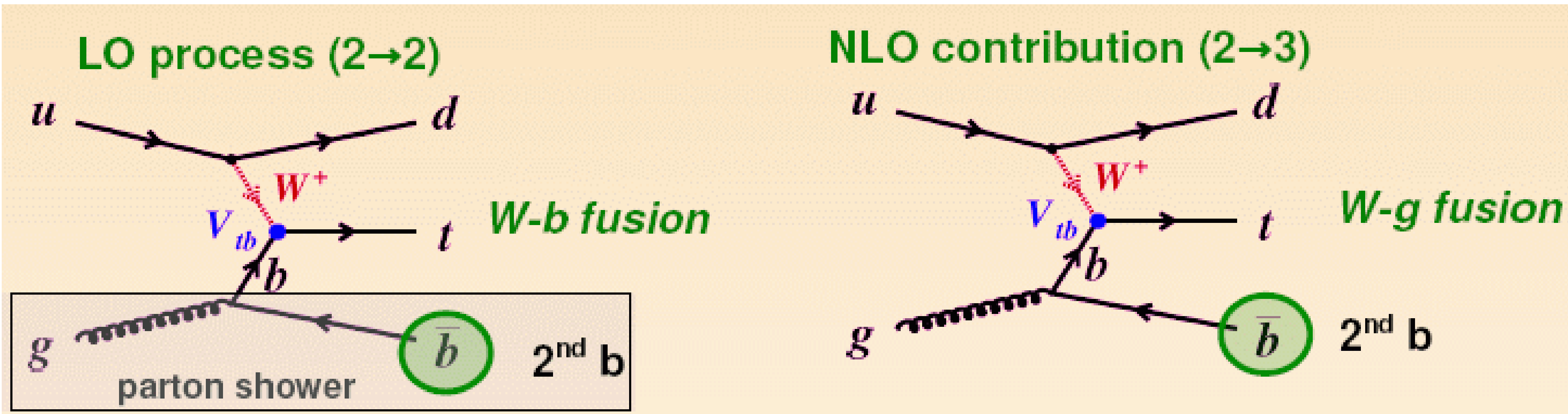


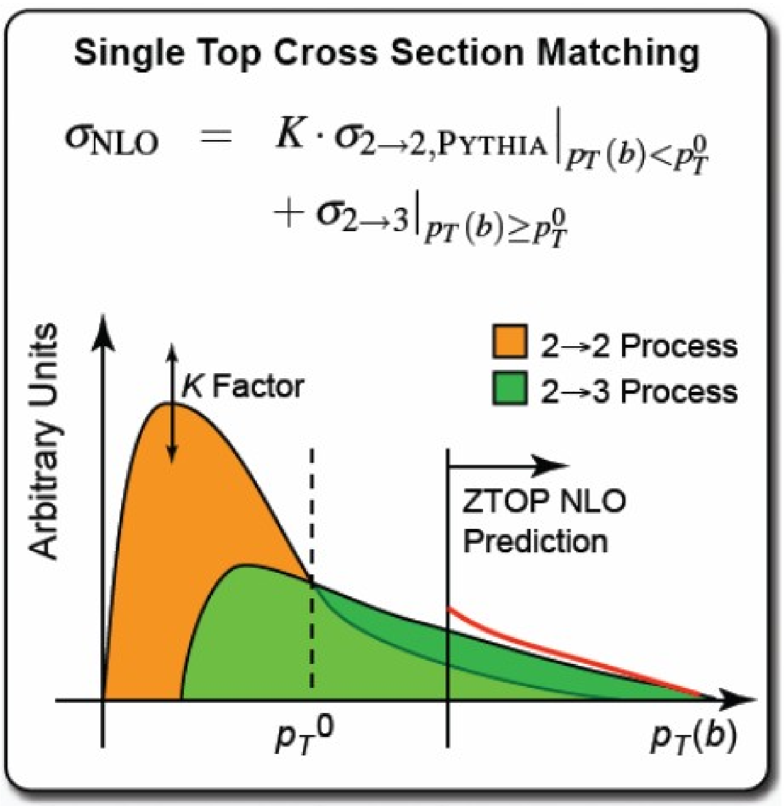
t channel (130 pb)

- Tevatron: recent 5σ observation in s+t channels (~1+2 pb)
  - tW negligible at 1.96 TeV
- LHC@10TeV: t chan. dominant
- s channel & tW are treated as backgrounds here

- Goals (increasing statistics):
  - **Confirmation of Tevatron**
  - Competitive constraint on $|V_{tb}|$
  - FCNC and anom. Wtb couplings; Other channels

Signal model: 2→2 / 2→3 matching





- Matching in p_T of the associated b
- Original idea: E.Boos, L.Dudko, V.Savrin, CMS NOTE 2000/065 (*SingleTop* gen.)
- Used in CDF (*MadGraph*), D0 (*SingleTop*)
- CMS implementation on top of *MadGraph*
  - Cross-validated with *SingleTop* and *MC@NLO* (AN2009/024)

Constraints on $|V_{ti}|$

- Measurements of V_{ub}, V_{cb} in the B sector + SM + 3x3 unitarity =

$$|V_{td}| \simeq 0.0069 - 0.0088$$

$$|V_{ts}| \simeq 0.0401 - 0.0418 \qquad \text{(at } 2\sigma \text{ level)}$$

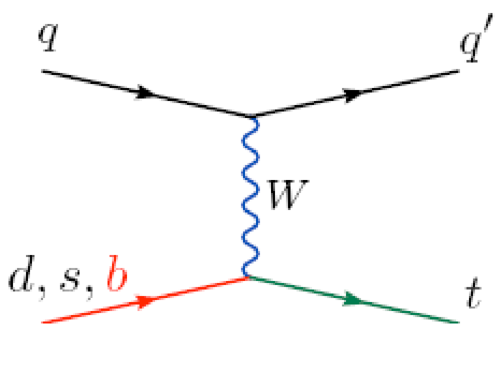
$$|V_{tb}| \simeq 0.9990 - 0.9992$$

- Measurements of ΔM_{Bs} and ΔM_{Bd} constrain $|V_{td}/V_{ts}|$
- Measuring R measures $|V_{tb}|$ only if 3x3 unitarity is assumed

$$R = \frac{\Gamma(t \to Wb)}{\Gamma(t \to Wq(=d,s,b))} = \frac{|V_{tb}|^2}{|V_{td}|^2 + |V_{ts}|^2 + |V_{tb}|^2}$$

Popular simplifying assumption: $|V_{ts,td}| \ll |V_{tb}|$ even if a 4th family exists
but D0 limit R> 0.79 only implies $|V_{tb}| > 1.9\sqrt{|V_{td}|^2+|V_{ts}|^2}$

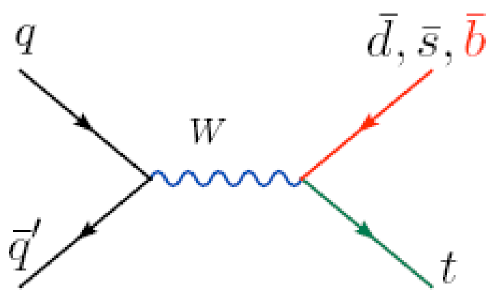
Direct constraints on $|V_{ti}|$, Tevatron



$$\sim |V_{td}|^2 \sigma_d^{\text{t-ch}} + |V_{ts}|^2 \sigma_s^{\text{t-ch}} + |V_{tb}|^2 \sigma_b^{\text{t-ch}}$$

Enhancement due to large *d* and *s* densities

(By the way: for the same reason, this is a good place to look for FCNC (*u* density))



$$\sim (|V_{td}|^2 + |V_{ts}|^2 + |V_{tb}|^2)\sigma^{\text{s-ch}}$$

Signal becomes similar to t-channel (only 1 *b*-jet)

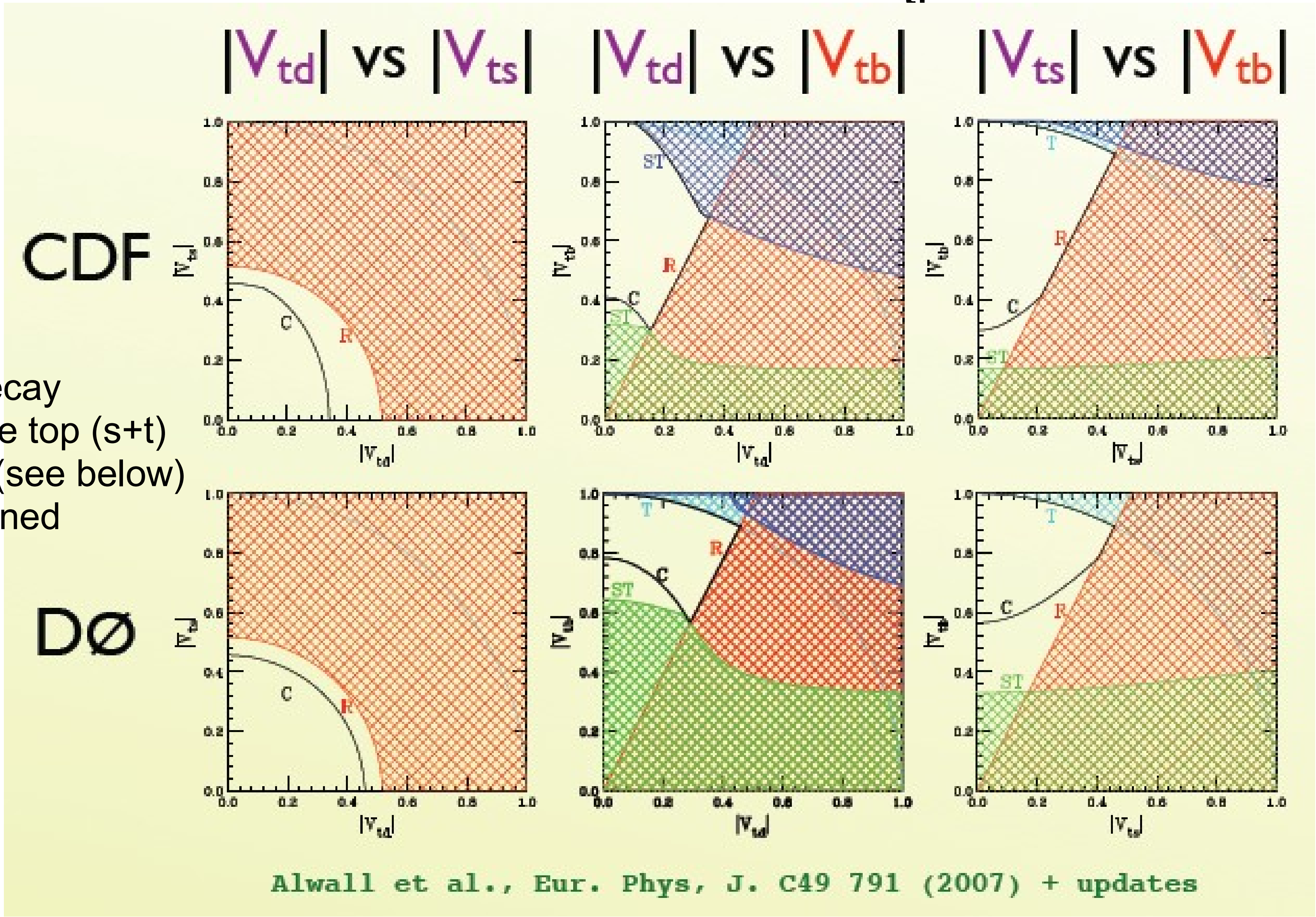
Very simplified meta-analysis of Tevatron results (ignoring differences in kinematics/topology – we would need access to the ntuples to do better):

$$\sigma_{1b\text{-tag}} = R\left\{\sum_{i=b,s,d} |V_{ti}|^2 \sigma_i^{t-ch} + 2(|V_{td}|^2 + |V_{ts}|^2)\sigma^{s-ch}\right\}$$

$$\sigma_{2b\text{-tag}} = R\,|V_{tb}|^2\,\sigma^{s-ch}$$

$$R = \frac{\Gamma(t \to Wb)}{\Gamma(t \to Wq(=d,s,b))} = \frac{|V_{tb}|^2}{|V_{td}|^2 + |V_{ts}|^2 + |V_{tb}|^2}$$

Direct constraints on $|V_{ti}|$, Tevatron



R: top decay
ST: single top (s+t)
T: trivial (see below)
C: combined

Simplifying assumption: no other new physics apart from new quarks
→ trivial constraint from Pythagoras' theorem

Direct constraints on $|V_{ti}|$, LHC

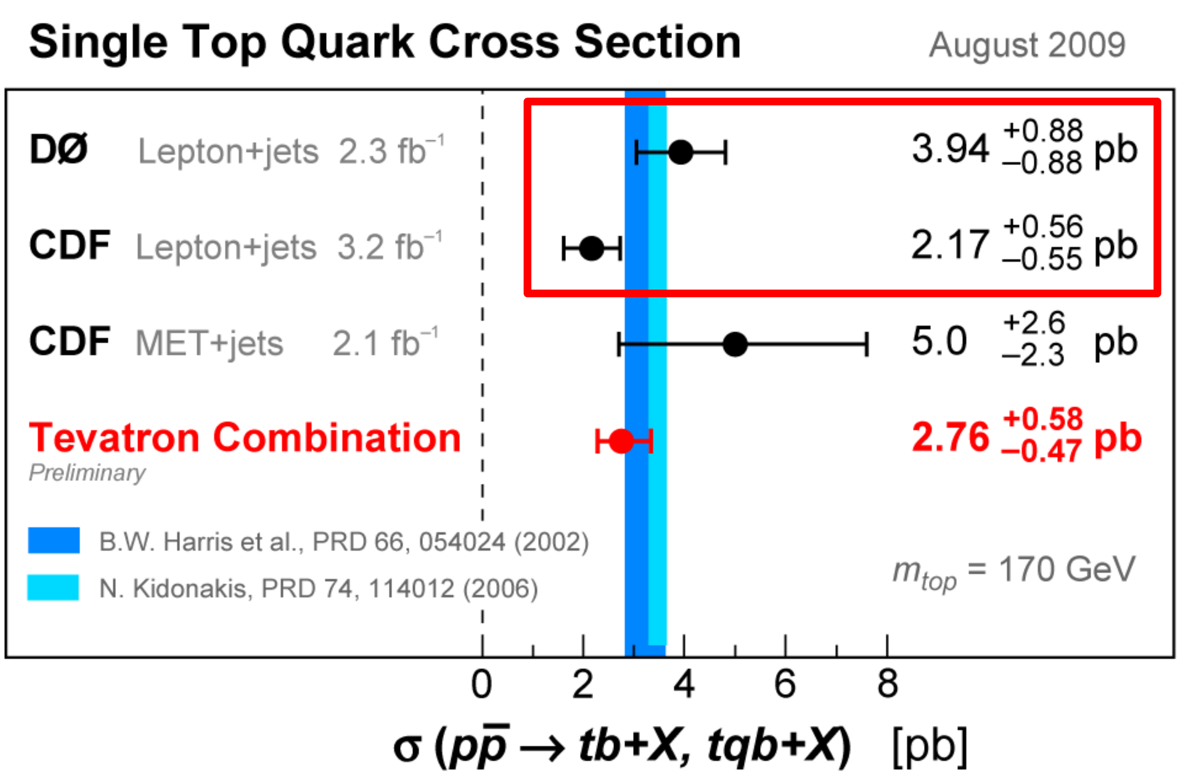
$$\sigma_{1b\text{-tag}} = R\left\{\sum_{i=b,s,d} |V_{ti}|^2 \sigma_i^{t-\mathrm{ch}} + 2(|V_{td}|^2 + |V_{ts}|^2)\sigma^{s-\mathrm{ch}}\right\}$$

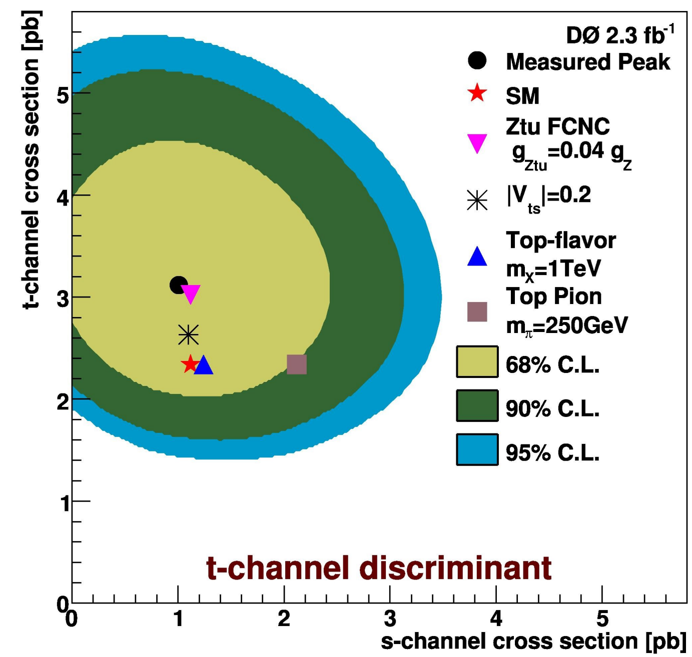
$$\sigma_{2b\text{-tag}} = R\,|V_{tb}|^2\,\sigma^{s-\mathrm{ch}}$$

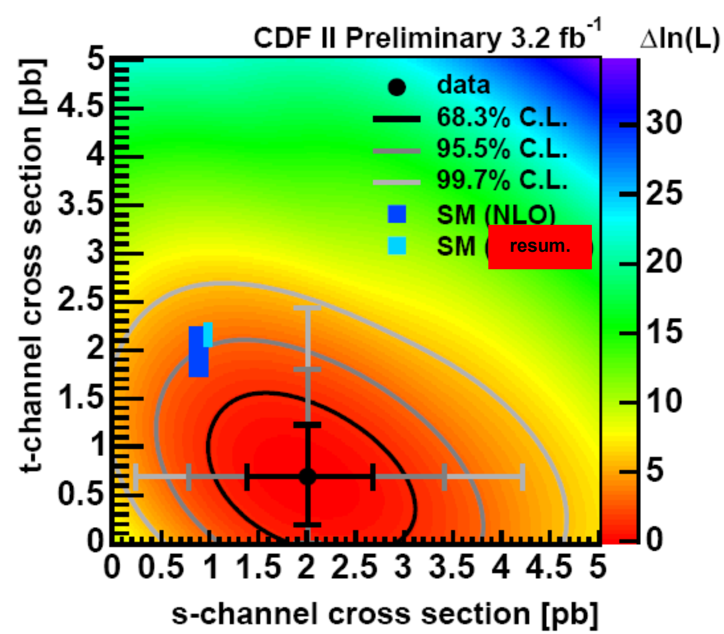
$$R = \frac{\Gamma(t \to Wb)}{\Gamma(t \to Wq(=d,s,b))} = \frac{|V_{tb}|^2}{|V_{td}|^2 + |V_{ts}|^2 + |V_{tb}|^2}$$

- At LHC, $\sigma^{s\text{-ch}} \ll \sigma^{t\text{-ch}}$
  - The 3rd channel, tW, is non negligible now; but it's “1b” too
- 2 measurements (σ from single top and R from ttbar) for 3 unknowns: top-only constraint of the entire 3rd row impossible
  - but we can use $|V_{td}/V_{ts}|$ from $\Delta M_{Bd,s}$ (which agrees with SM)
- Expected precision on R @ 10 TeV:
  - ±2%(stat)±9%(syst) with 250/pb, CMS-PAS-TOP-09-001, e+μ+2j channel only; systematics expected to halve with 10x data

The Tevatron "tension"



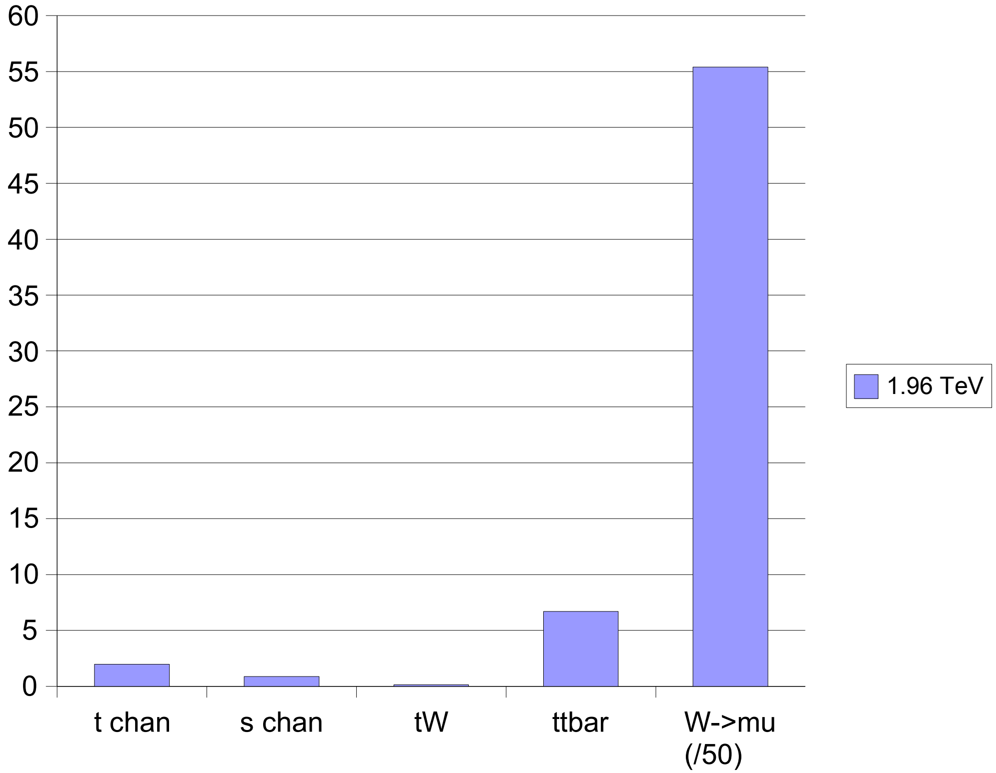


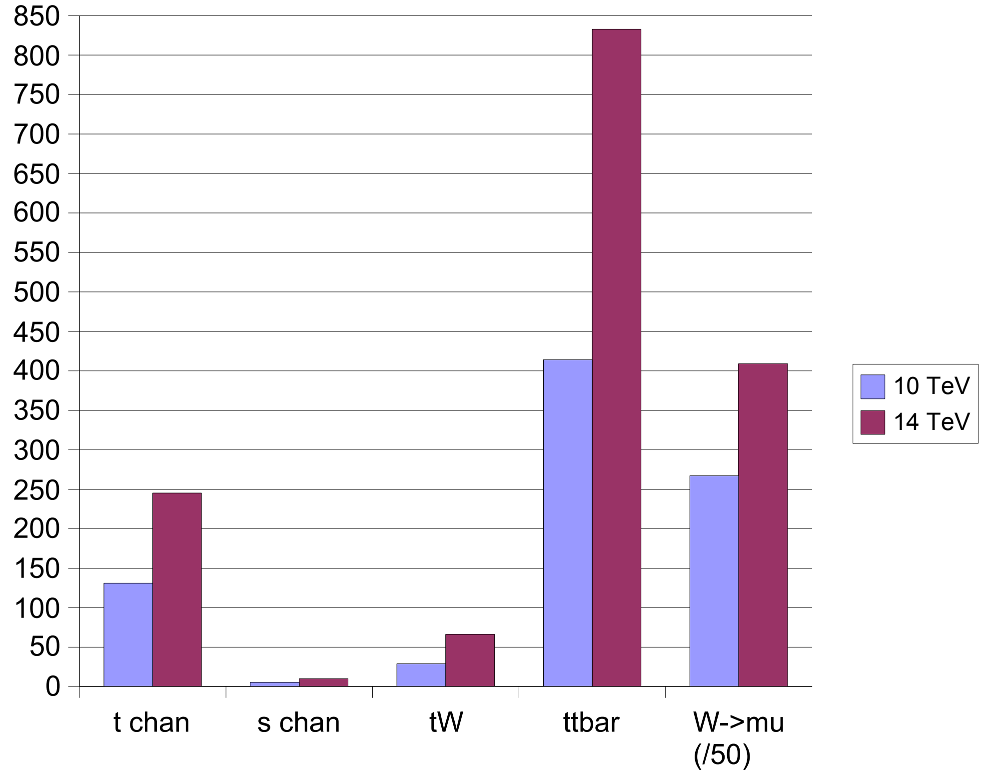


D0-CDF difference seems mostly attributable to the *t* channel alone (deficit in CDF, excess in D0)
→ blame the *b* pdf?

From Tevatron to LHC

Cross sections (pb):





LHC analyses will be quite different from Tevatron:
W+jets will be less of a concern than ttbar

LHC start-up plans

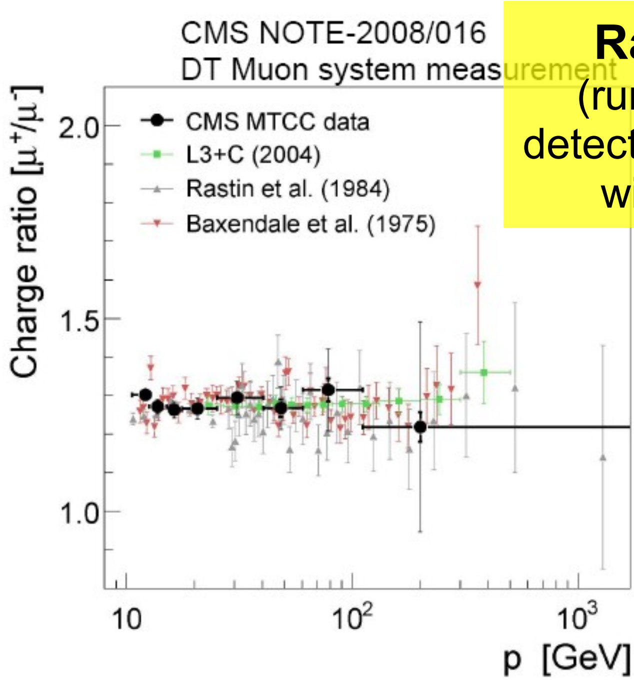
http://lhc-commissioning.web.cern.ch/lhc-commissioning/luminosity/09-10-lumi-estimate.htm

Month	Comment	Turn around time	Energy [TeV]	Max number bunches	Protons/Bunch	% nom. intensity	Min beta*	Peak Luminosity $cm^{-2}s^{-1}$	Integrated Luminosity	events/X
1	Beam commissioning								First collisions	
2	Pilot physics, partial squeeze, gentle increase in bunch intensity, avaialbility low	Long	3.5	43	3×10^{10}		4 m	8.6×10^{29}	100 - 200 nb^{-1}	
3		5	3.5	43	5×10^{10}		4 m	2.4×10^{30}	~ 1 pb^{-1}	
4		5	3.5	156	5×10^{10}	2.5	2 m	1.7×10^{31}	~9 pb^{-1}	
5a	No crossing angle - could at this stage push intensity see 5b	5	3.5	156	7×10^{10}	3.4	2 m	3.4×10^{31}	~18 pb^{-1}	0.8
5b	No crossing angle - squeezing to beta* = 1m at this stage would double these lumi numbers (and the pile-up)	5	3.5	156	10×10^{10}	4.8	2 m	6.9×10^{31}	~36 pb^{-1}	1.6
6	Possible shift to higher energy - would anticipate ~4 weeks to reestablish physics follow by a fairly gentle increase back up in intensity.	Would aim to first provide a period of physics at the higher energy (4.5 TeV, say) without crossing angle, this could be followed by a move to 50 ns with a limited number of bunches. Note that the total intensity limit will go down with the move to higher energy.								
7	4 - 5 TeV (5 TeV luminosity quoted - doesn't make too much difference). No crossing angle.	5	4 -5	156	7×10^{10}	3.4	2 m	4.9×10^{31}	~26 pb^{-1}	
8	50 ns - nominal crossing angle - aperture restricts squeezing further - note limited complement of bunches.	5	4 -5	144	7×10^{10}	3.1	2 m	4.4×10^{31}	~23 pb^{-1}	
9	50 ns	5	4 -5	288	7×10^{10}	6.2	2 m	8.8×10^{31}	~46 pb^{-1}	
10	50 ns*	5	4 -5	432	7×10^{10}	9.4	2 m	1.3×10^{32}	~69 pb^{-1}	
(11)	50 ns*	5	4 -5	432	9×10^{10}	11.5*	2 m	2.1×10^{32}	~110 pb^{-1}	

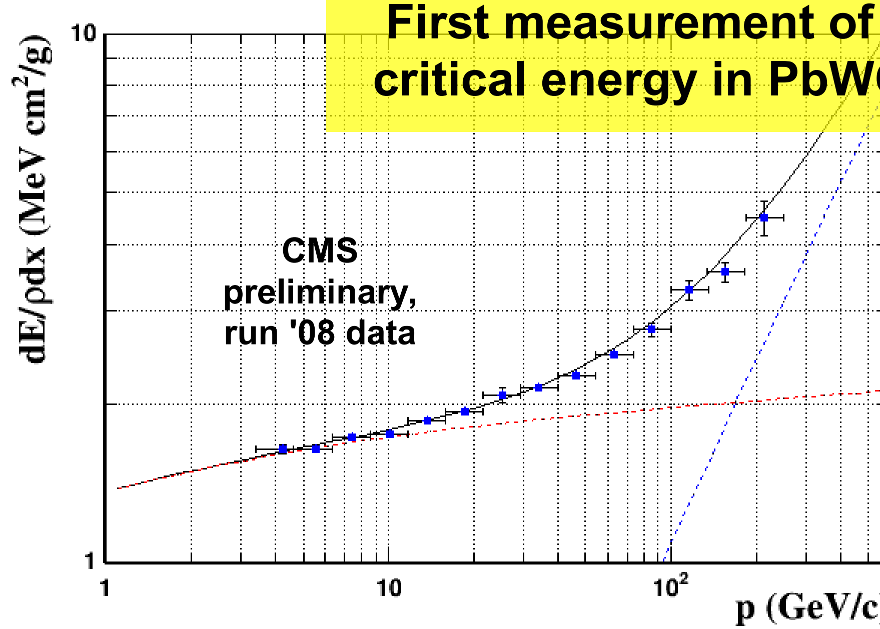
~60/pb @ 7 TeV, then 200-300/pb @ 8-10 TeV

Status of CMS

- Long cosmic runs in 2008 and 2009, very useful for commissioning
- Rarely (if ever) in the history of collider physics a detector, at the time of the first collisions, has been as ready as CMS and ATLAS
- In 2007 we estimated the alignment uncertainties realistically achievable at the time of 0, 10 and 100 /pb; current alignment with cosmics achieved, in the barrel, the goal of that 100/pb scenario...
- Even doing some real physics w/ cosmics:



Ratio μ^+/μ^- vs P
(run '06, incomplete detector; new paper soon with run '08 data)

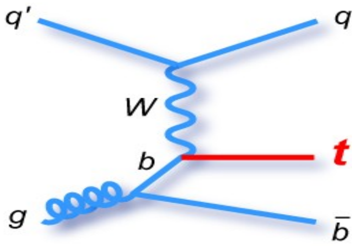


First measurement of μ critical energy in $PbWO_4$

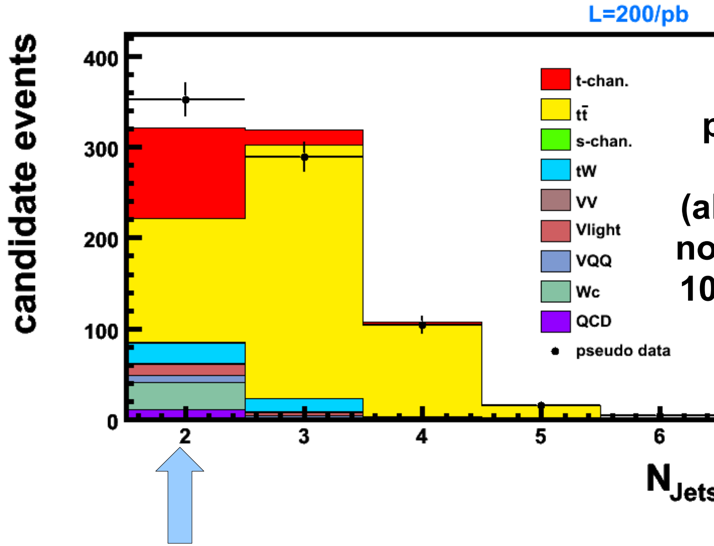
Our analysis at 10 TeV: event selection

Hunts Needle in a Haystack

Cuts overview



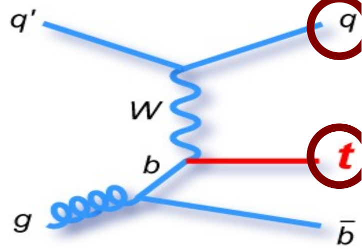
- **Single muon, di-lepton veto, isolated, far from any jet**
  - 1μ (p_T>20 GeV, |η|<2.1, plus some quality cuts)
  - 0e (p_T>20 GeV, |η|<2.4, plus tight identification cuts)
  - relIso = $p_T/(p_T$+tkIso+caloIso) >0.95
- **Two jets, far from the muon**
  - Iterative cone R=0.5 (not critical)
  - p_T>30 GeV, |η|<5
  - ΔR(μ,jets)>0.3 otherwise the event is discarded ("near-jet veto")
- **One b jet**
  - "Track counting" tagger
  - 1j passing a tight selection
  - 2nd jet: it must fail a loose selection
- **On-shell W boson (t→Wb)**
  - M_T>50 GeV

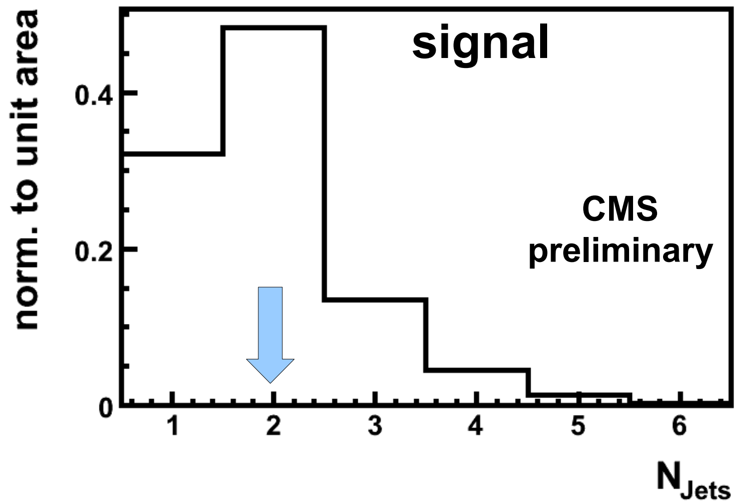


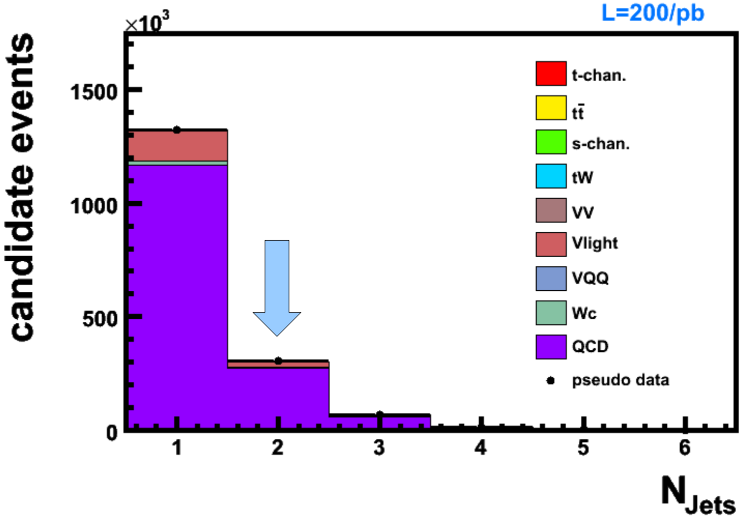
CMS preliminary

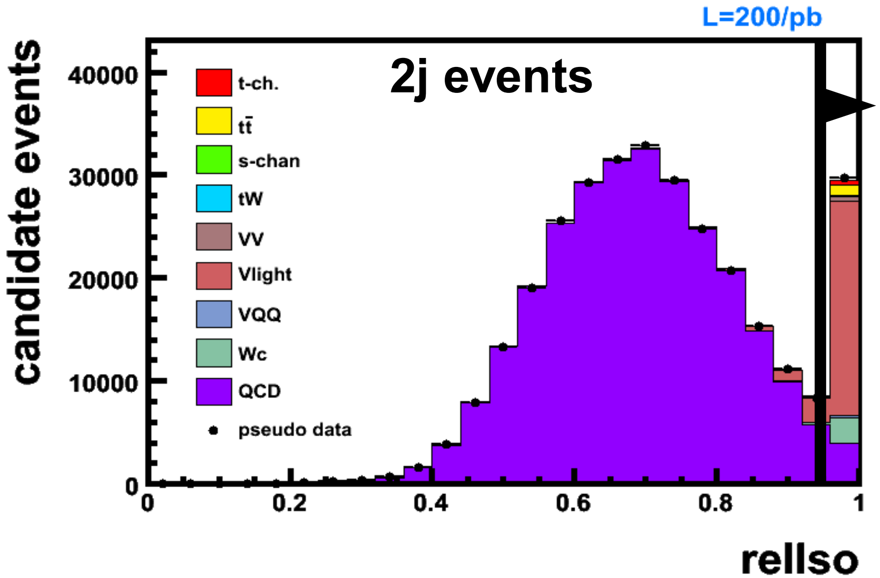
(all plots from now on are for 10 TeV unless specified)

After lepton and jet counting (1μ, 0e, 2j), the sample is still QCD-dominated





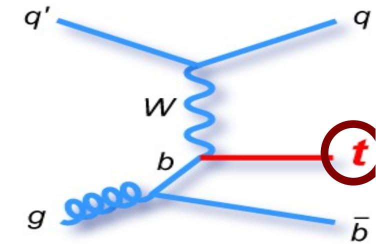


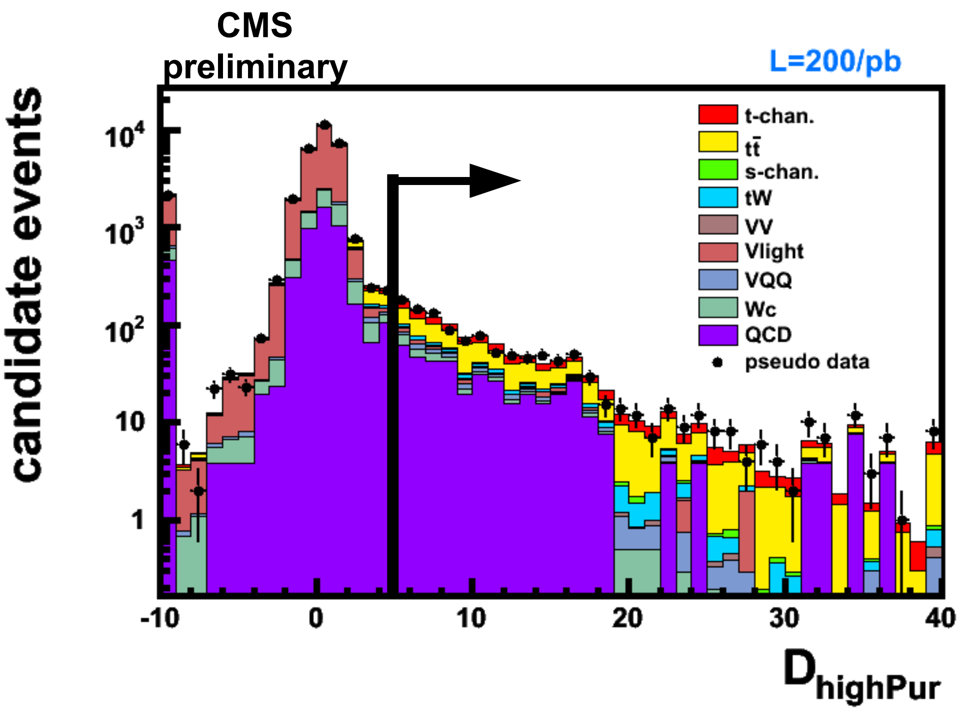


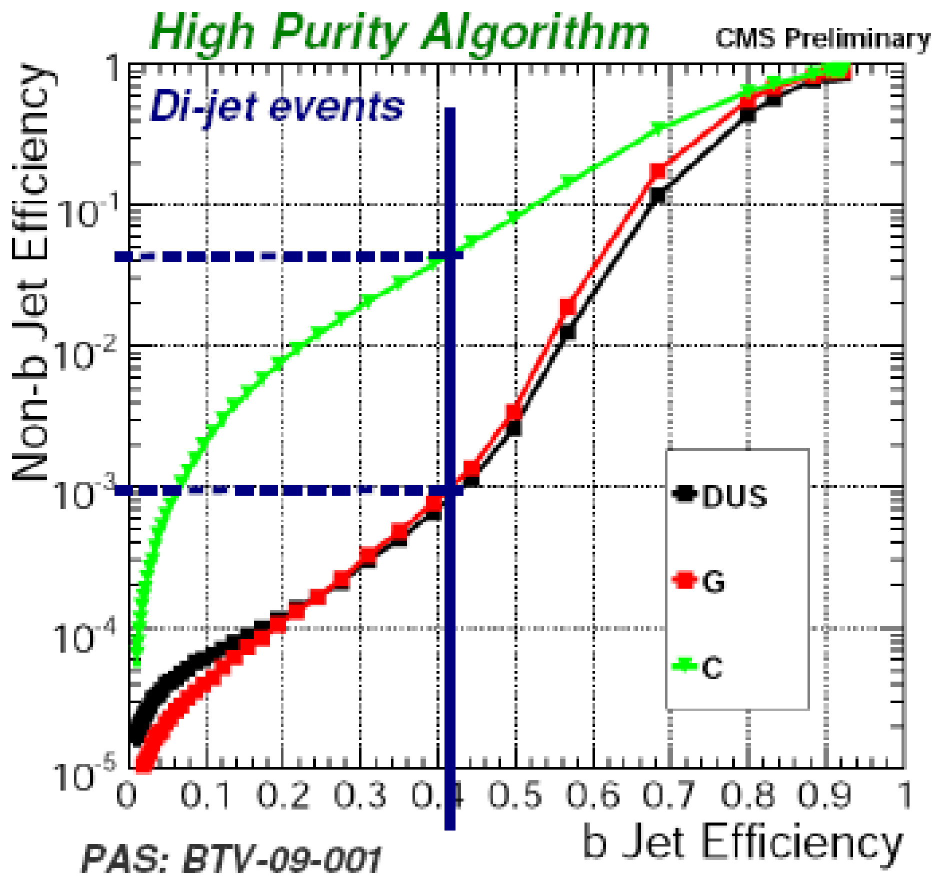
$$relIso = \frac{p_{T,\mu}}{p_{T,\mu} + tkIso + caloIso}$$

After the isolation request, it is W-dominated

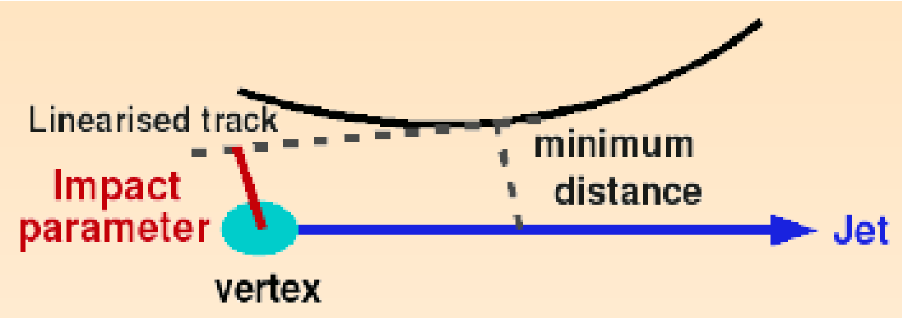
Getting rid of W+jets: tight b tagging



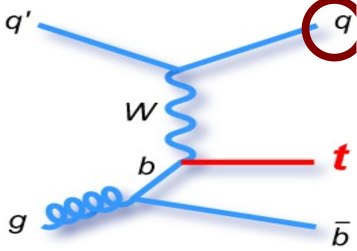


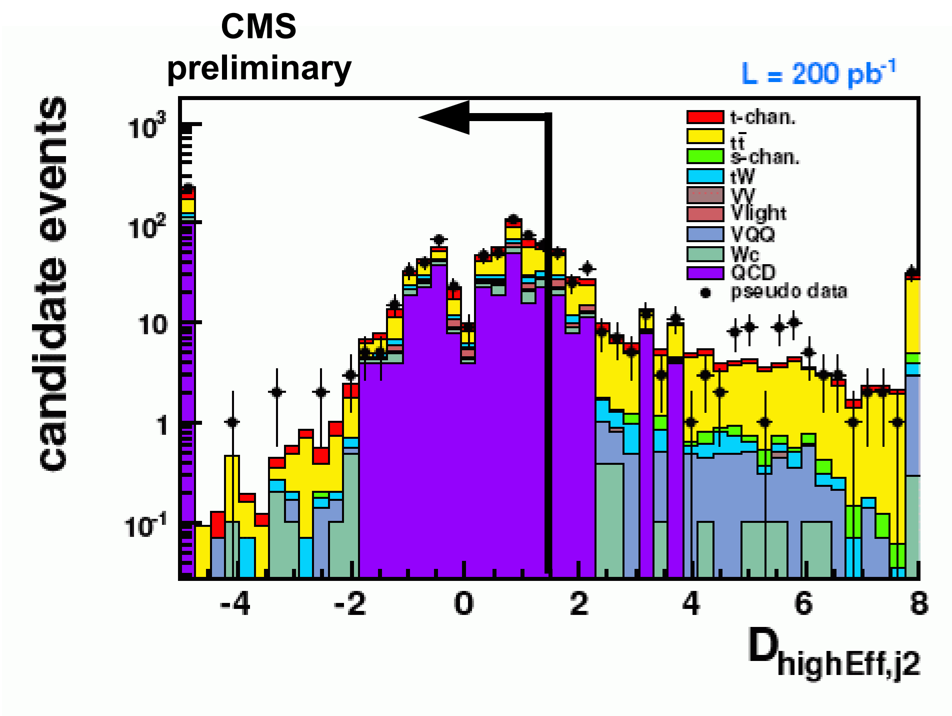


High-purity track-counting algorithm, i.e. 3rd track IP/σ_{IP} in the jet

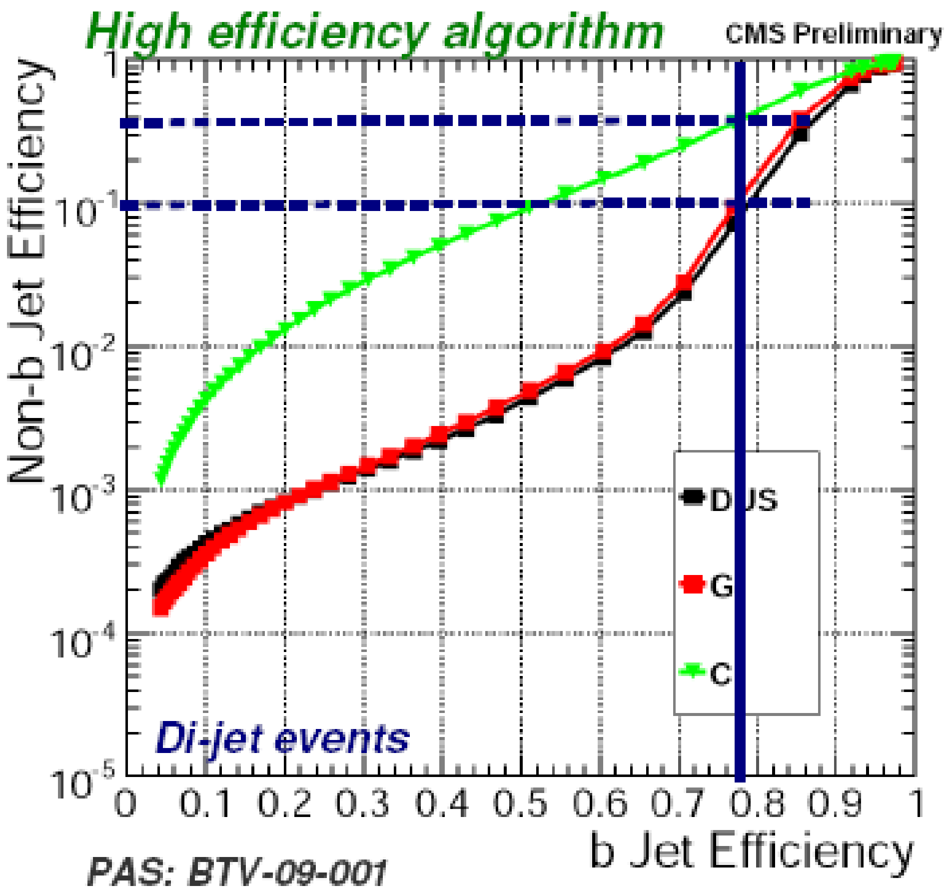


Signal has 1b only, most bkgs 2b: veto on 2nd b-jet, loose b tagging

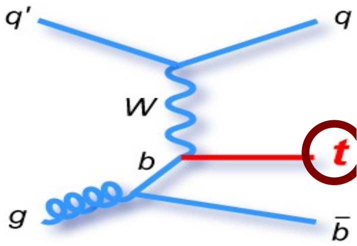


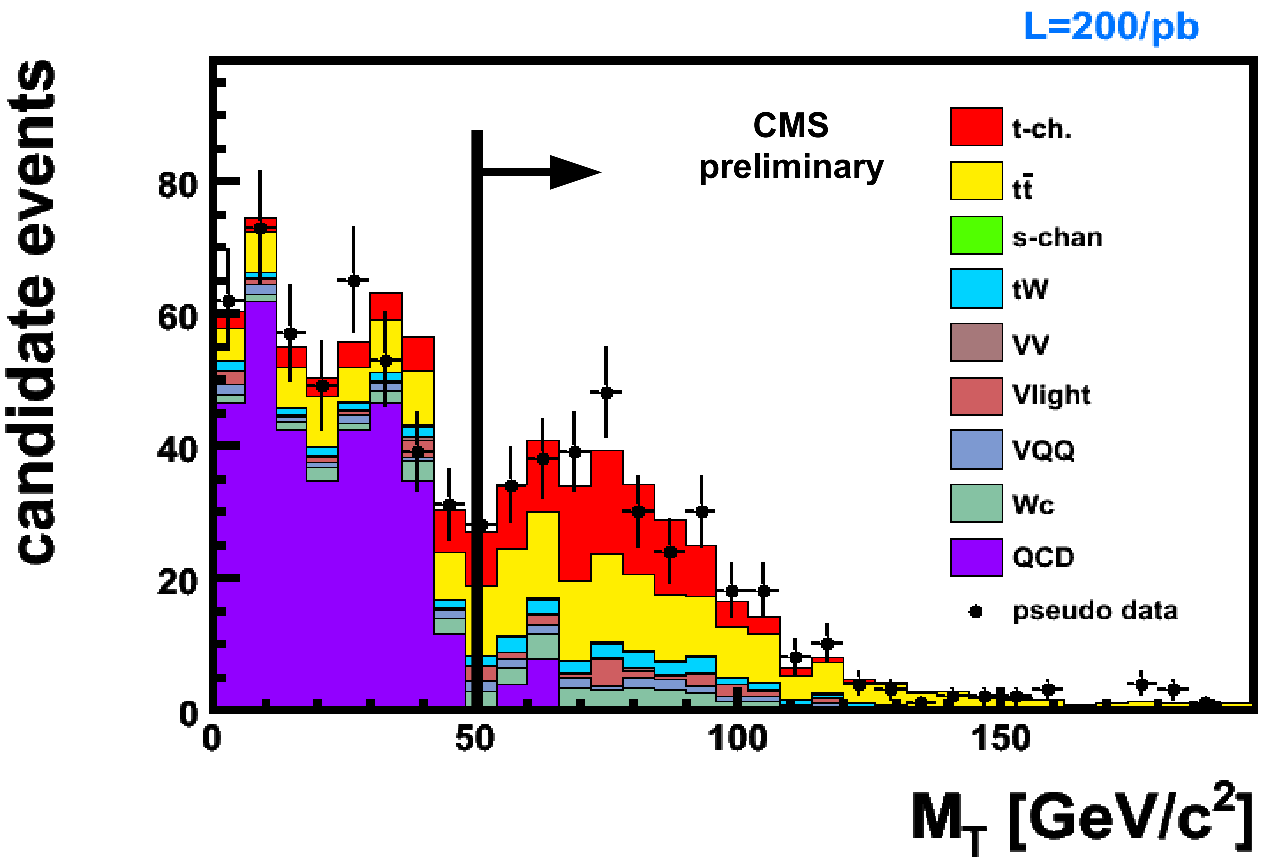


High-efficiency track-counting algorithm, i.e. 2rd track IP/σ_{IP} in the jet



Invariant transverse mass (M_T): QCD has no Jacobian peak

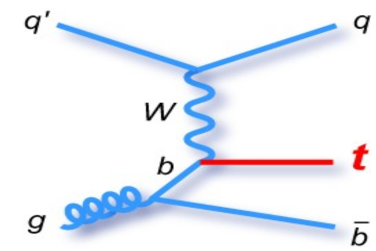




For an on-shell W boson:

$$M_T = \sqrt{(p_{T,\mu} + p_{T,\nu})^2 - (p_{x,\mu} + p_{x,\nu})^2 - (p_{y,\mu} + p_{y,\nu})^2}$$

Expected yield in 200/pb @ 10 TeV



Process	$\sigma \times$ BR[pb]	L $[fb^{-1}]$	N_{evt} in 200 pb^{-1}
single top, t channel ($W \to l\nu, l = e, \mu, \tau$)	42.9 (NLO)	6.6	102±1.8
single top, s channel ($W \to l\nu, l = e, \mu, \tau$)	1.6 (NLO)	7.5	1.8±0.2
single top, tW	29 (NLO)	5.8	22.3± 0.9
$t\bar{t}$	414 (NLO+NLL)	2.2	136.0±3.5
QCD multi-jet (μ-enriched)	121675(LO)	0.05	12±6.7
Wc ($W \to l\nu, l = e, \mu, \tau$)	1 490 (LO)	2.0	29±1.7
$Wb\bar{b}$ ($W \to l\nu, l = e, \mu, \tau$)	54.2 (LO)	2.9	8.0±0.7
$Wc\bar{c}$ ($W \to l\nu, l = e, \mu, \tau$)	118.8 (LO)	4.5	1.2±0.2
$W+$ light partons ($W \to l\nu, l = e, \mu, \tau$)	40 000 (LO)	0.24	12±2.6
$Zb\bar{b}$ ($Z \to ll, l = e, \mu, \tau$)	44.4 (LO)	3.5	2.7±0.4
$Zc\bar{c}$ ($Z \to ll, l = e, \mu, \tau$)	71.7 (LO)	5.0	0.2±0.1
$Z+$ light partons ($Z \to ll, l = e, \mu, \tau$)	3 700 (LO)	0.33	2±1.2
WW	74 (LO)	2.8	0.9±0.3
WZ	32 (LO)	7.4	1.2±0.2
ZZ	10.5 (LO)	19.0	0.17±0.04
Total Background			229±8.4

S/B=0.45

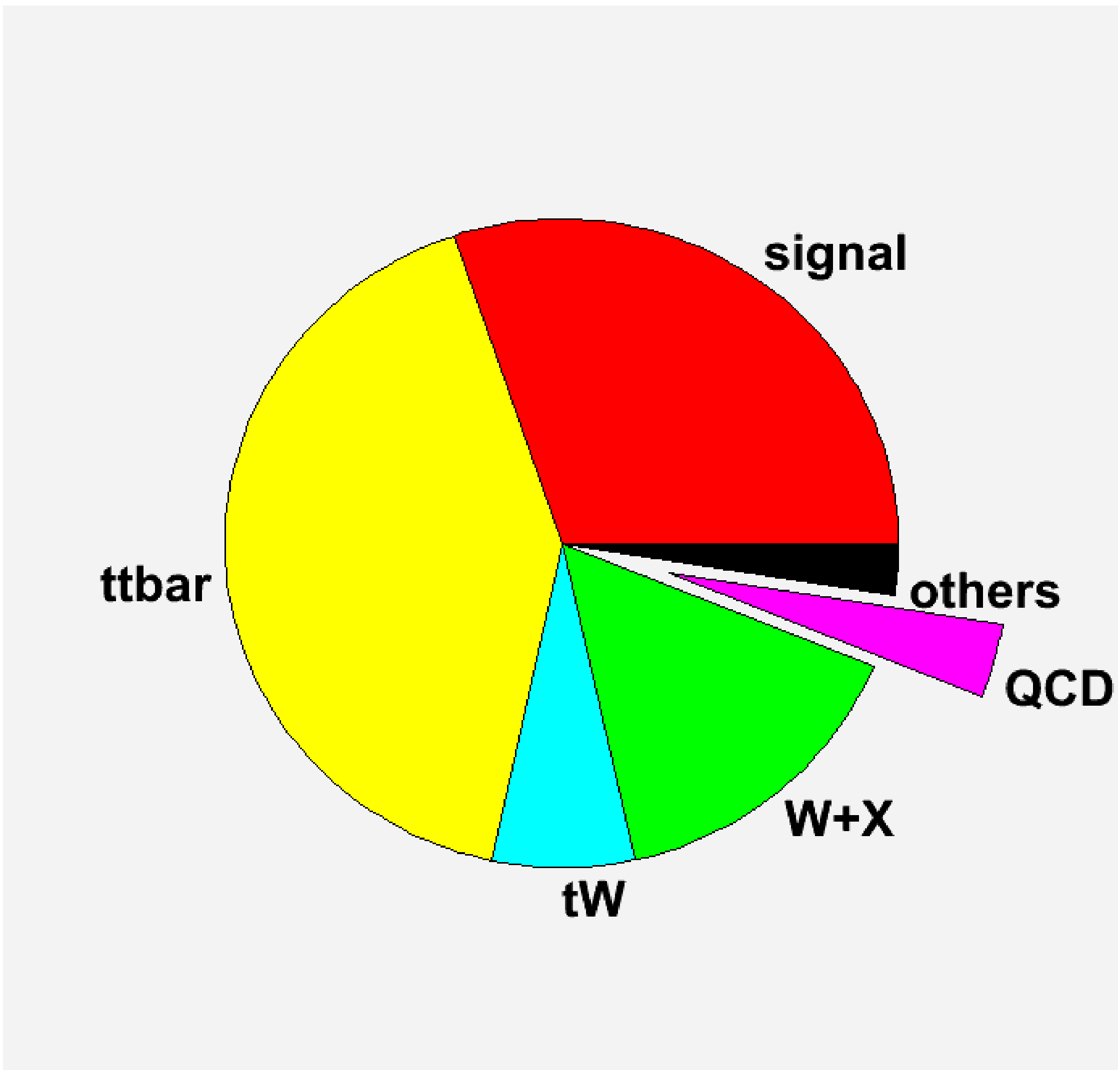
S/√B=6.7

S/√S+B=5.6

A counting experiment would require a level of knowledge of B that seems unrealistic in this scenario.

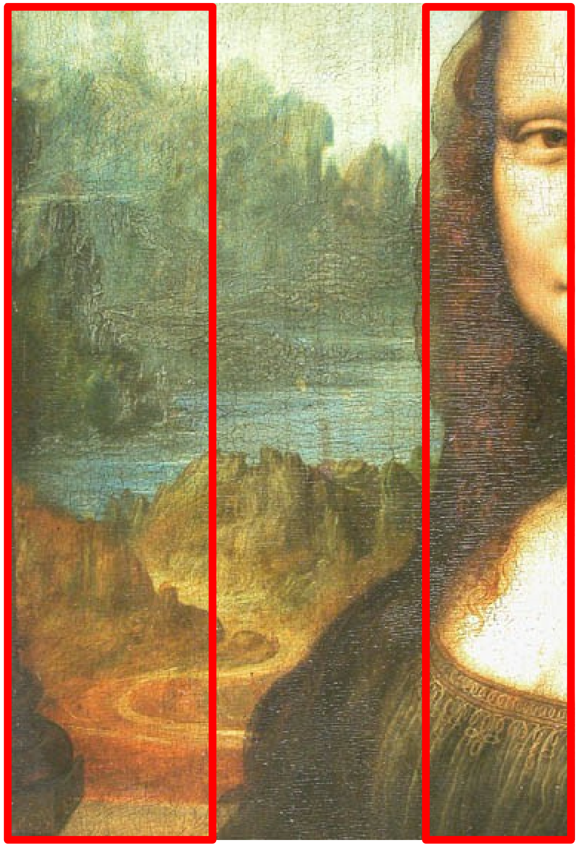
We need a method with minimal need of assumptions about B

Multi-jet QCD estimation



QCD is not the dominant background (in MC!), but it's the least predictable one (not only in the abundance, but also in the shapes of the discriminating variables that I will introduce later: how to trust MC? Any QCD event able to pass our selection is a very atypical one...)

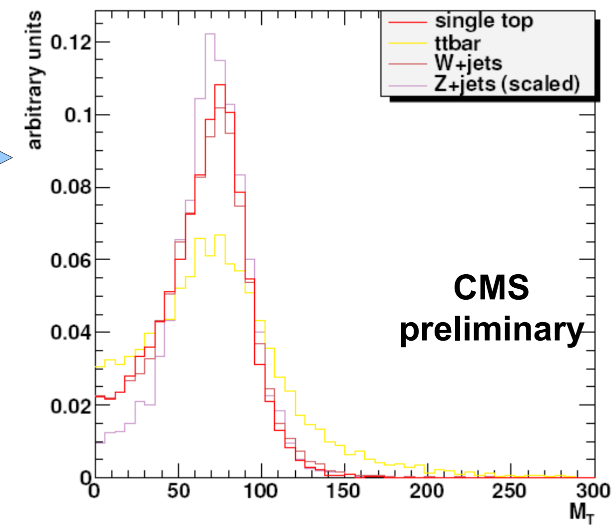
Multi-jet QCD estimation



1) See how a background-enriched region looks like

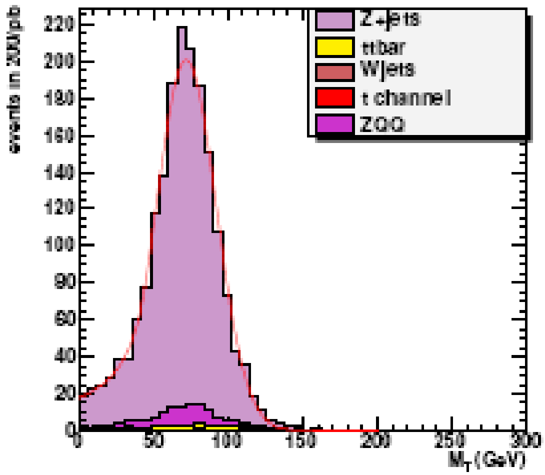
2) Use this information to guess what's behind the signal in the analysis region

Strategy



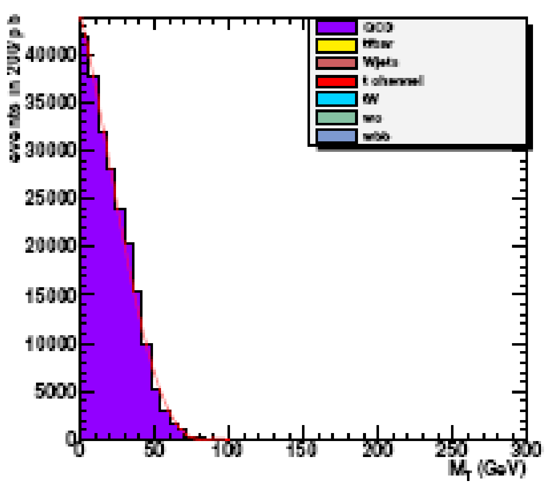
- Signal/tW/tt/WX roughly similar in $\mathbf{M_T}$
- After full sel., fit to $F(M_T)=aS(M_T)+bB(M_T)$
- Minimize model assumptions:
  - shapes $S(M_T)$, $B(M_T)$ are both taken from **control samples:**
  - QCD-enriched: no b tag cut, relIso<0.8, all the rest the same
  - Z-enriched: 2μ, 2j, no b tag cut, $76<M_{\mu\mu}<106$ GeV
    - Muon momenta rescaled by M_W/M_Z
    - A μ, randomly chosen, is treated as a ν (summed to MET)
    - Purity very high, and M_T shape resembles signal enough
  - Alternative $S(M_T)$ models: MC truth, or W-enriched (no btag)

Control samples => prediction



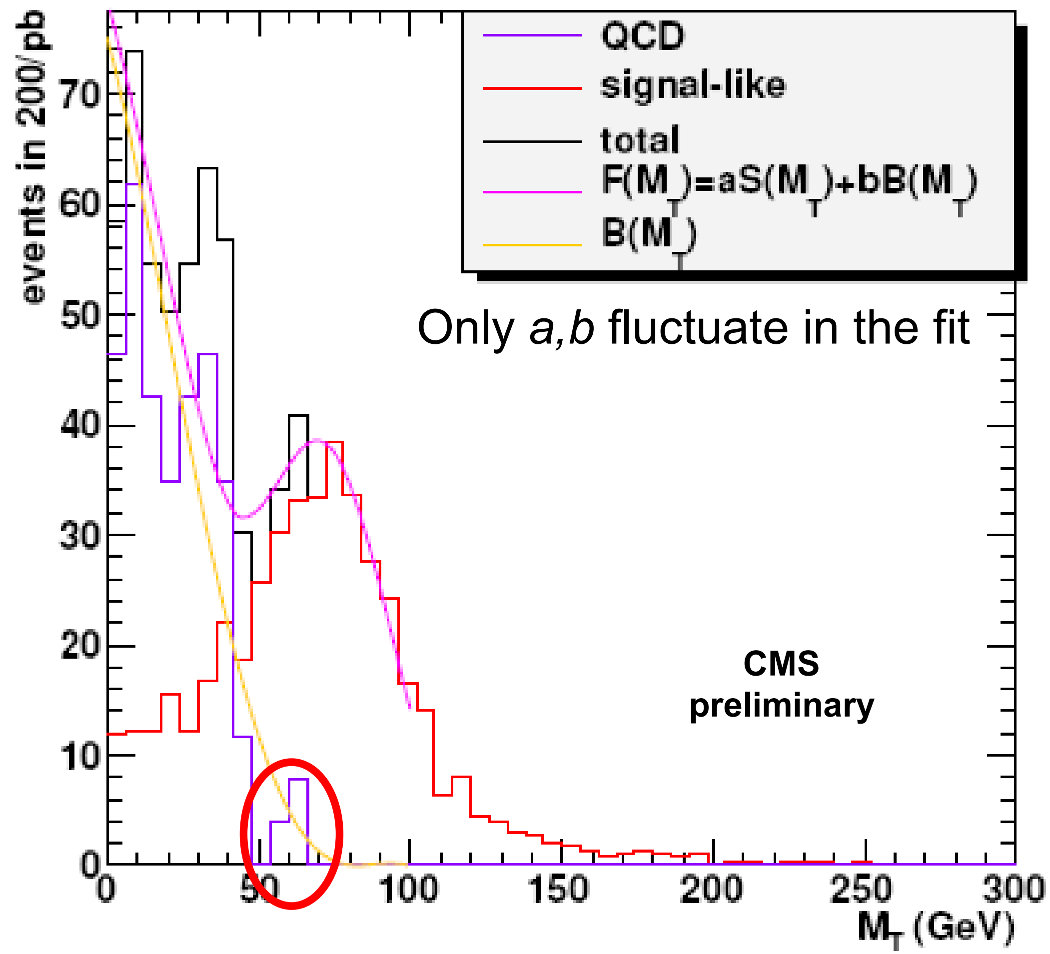
S(M_T): Z-enriched sample

fitted to a “Crystal ball” function



B(M_T): QCD-enriched sample

fitted to a polynomial of rank 4



Result: 22.0 QCD events predicted for M_T>50 GeV, versus 12±7 actual. Stat.error from the fit: 6%. S(M_T) from W-enriched: 15.1; pure MC: 19.7

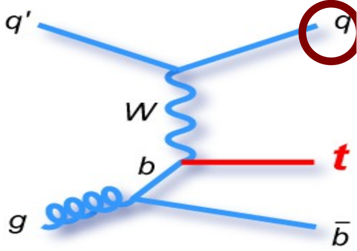
Is there single top in our sample?

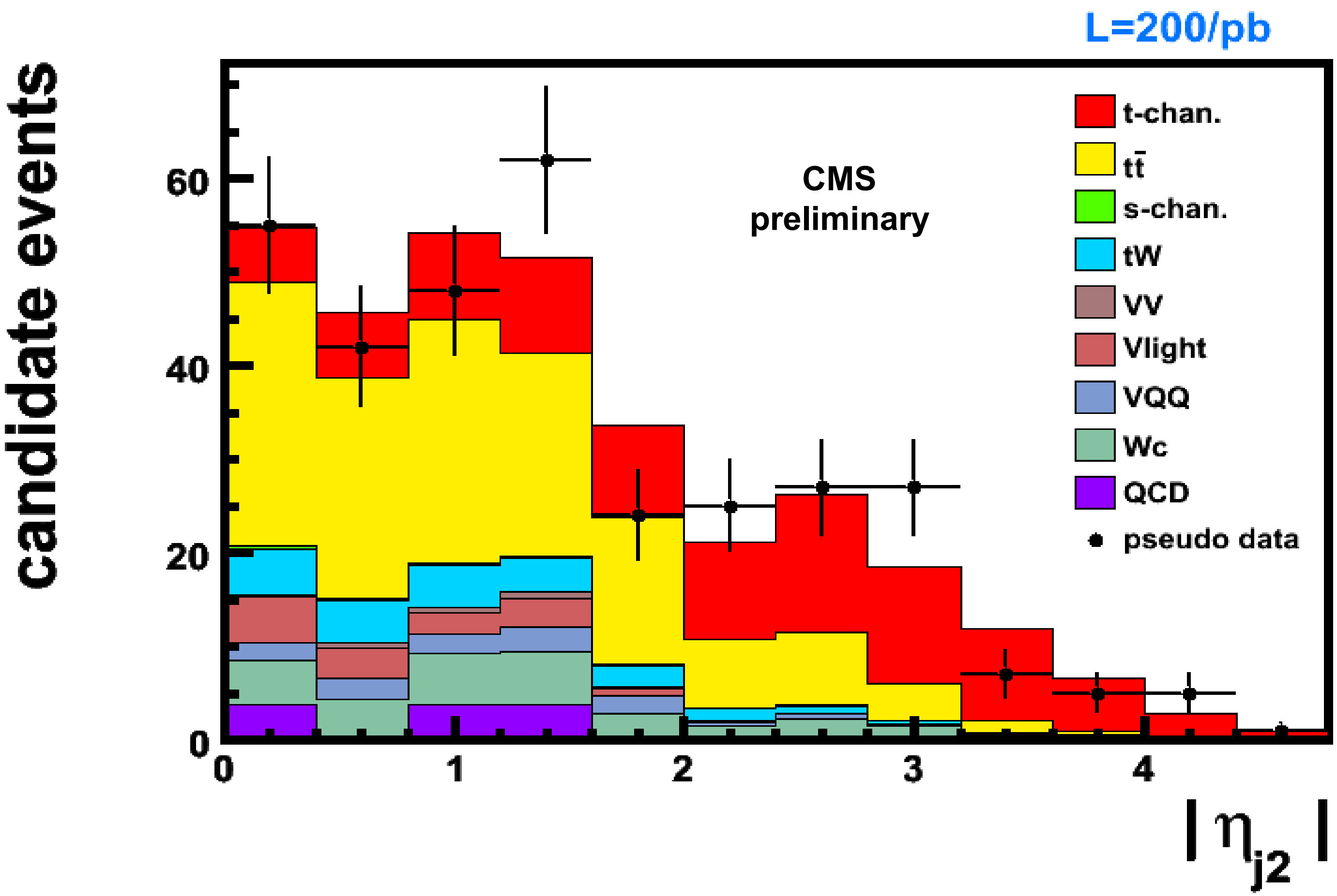
Some observables, for XXX pb^{-1} of data passing full selection

Our sausage machine could be a simple "cut and count", a powerful MVA technique, or whatever; and dozens of discriminating variables could be its input. Our choices were driven by the specificities of an "early data" scenario.

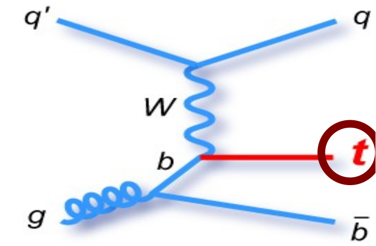
Background-only hypothesis is excluded at Nσ level (eventually, cross section measurement)

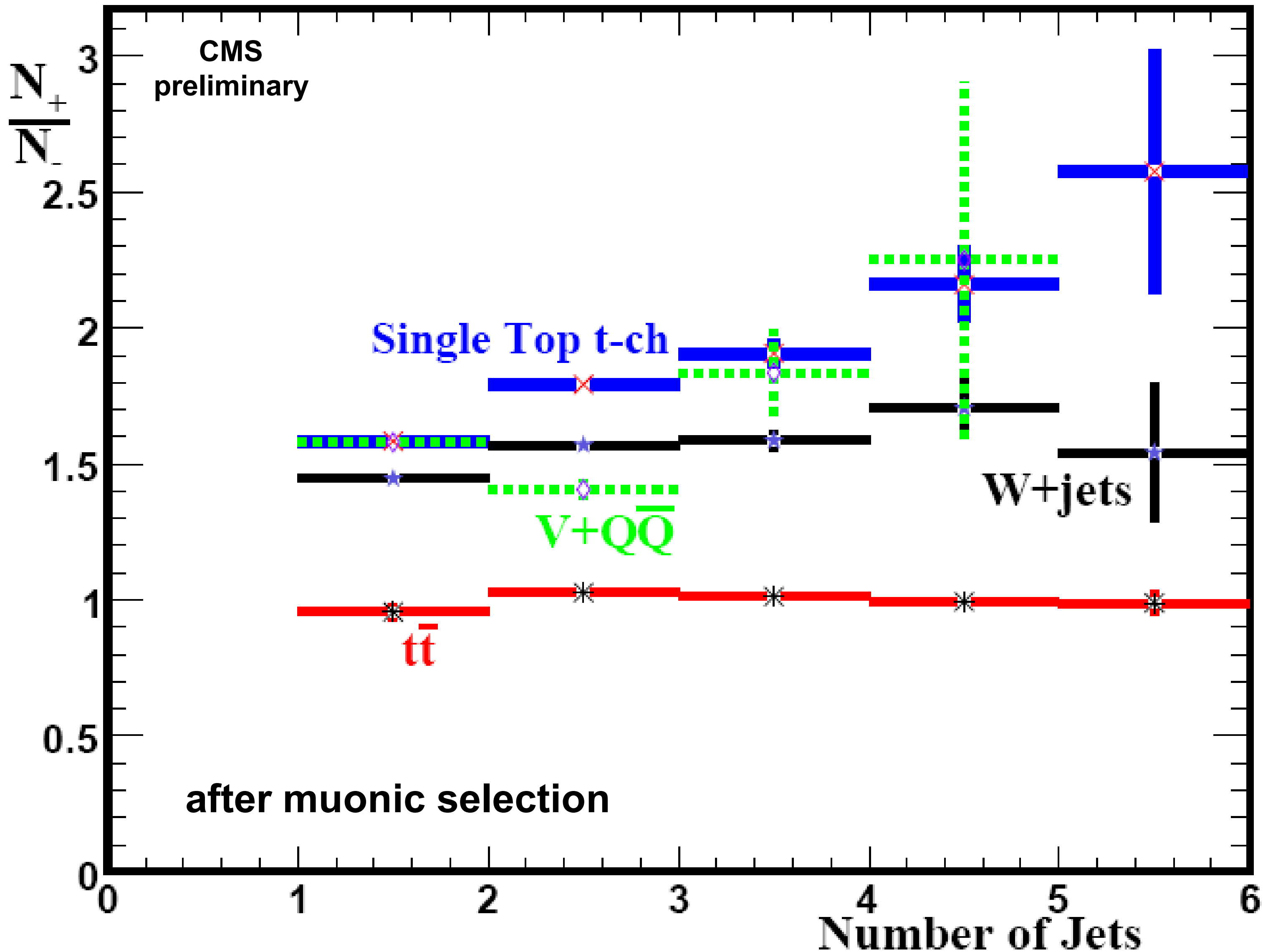
Smoking gun #1: "recoil quark"



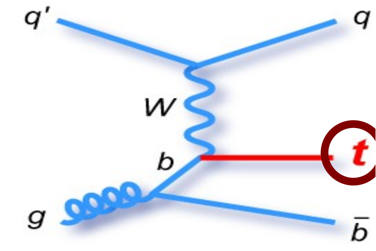


Smoking gun #2: charge asymmetry

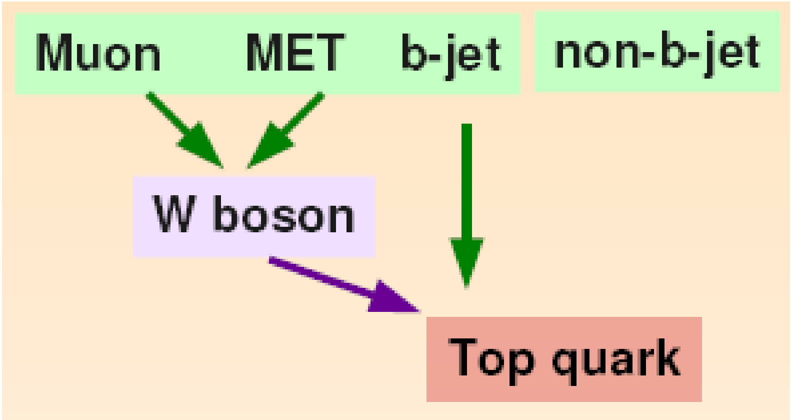


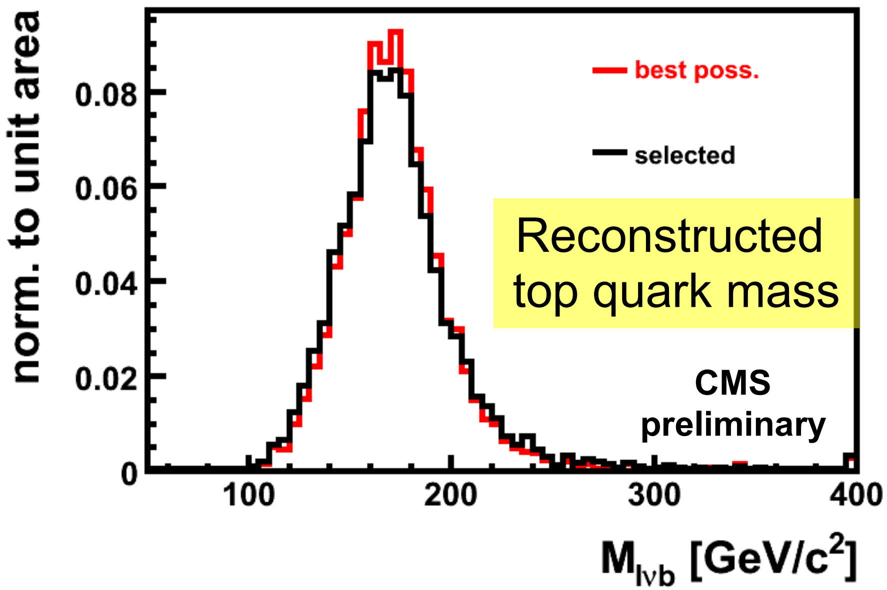


Top quark reconstruction



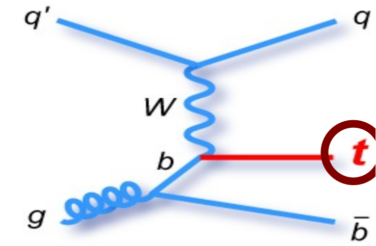
- W boson reconstruction:
  - W mass constraint
  - 2nd order equation in $P_{z,\nu}$
  - Complex solutions (36% of sel.evts.)
    - Forcing $M_T = M_W \rightarrow \mathrm{Img}(P_{z,\nu}) = 0$
  - Two real solutions (64% of sel.evts.)
    - Pick the one with smallest $P_{z,\nu}$
- Pairing with a b:
  - Just take the b-tagged jet
  - Correct in 92.2% of selected events
  - The associated b accounts only for 4.0%

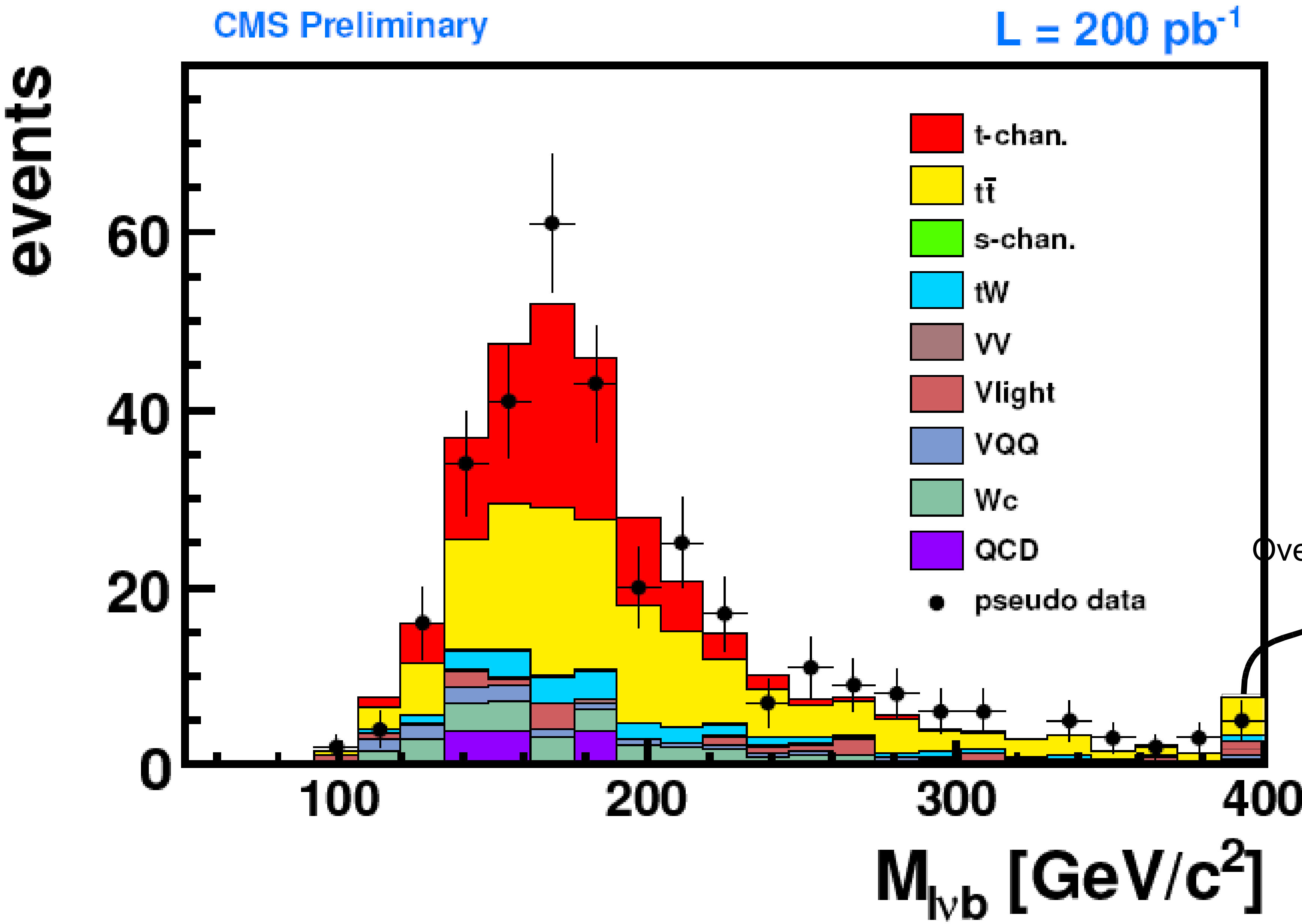




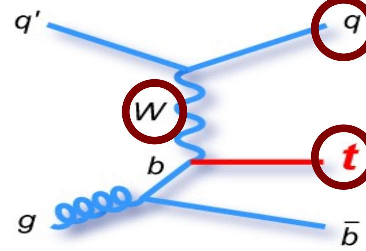
"Best possible" means minimal distance (ΔR) of reco objects to partons

Smoking gun #3: a peak at the right mass



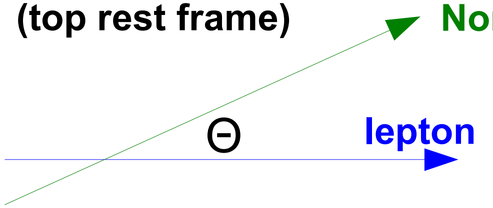


Smoking gun #4: this top quark is polarized



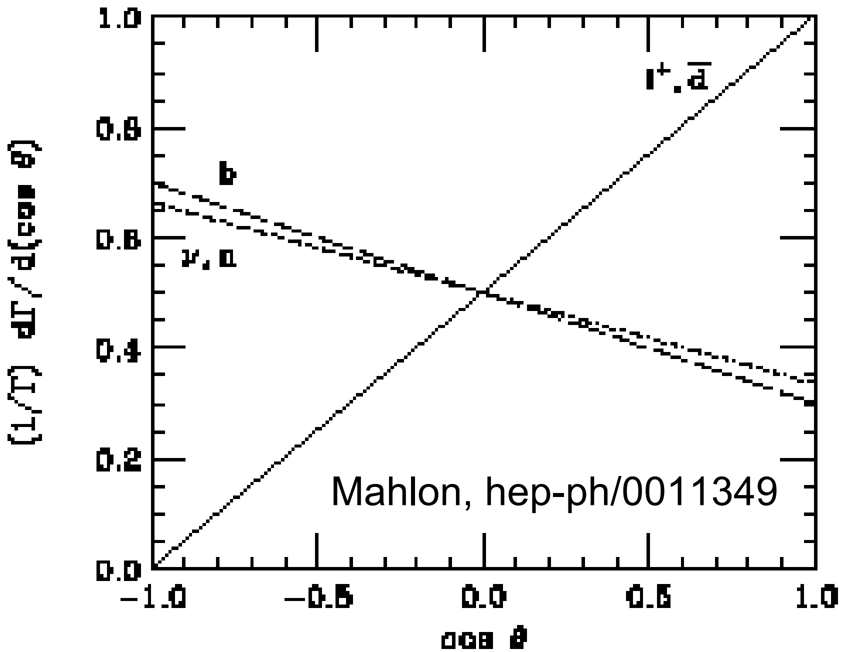
Most characteristic feature of single top events, stemming from the V-A nature of the Wtb coupling; **propagated to decay products**

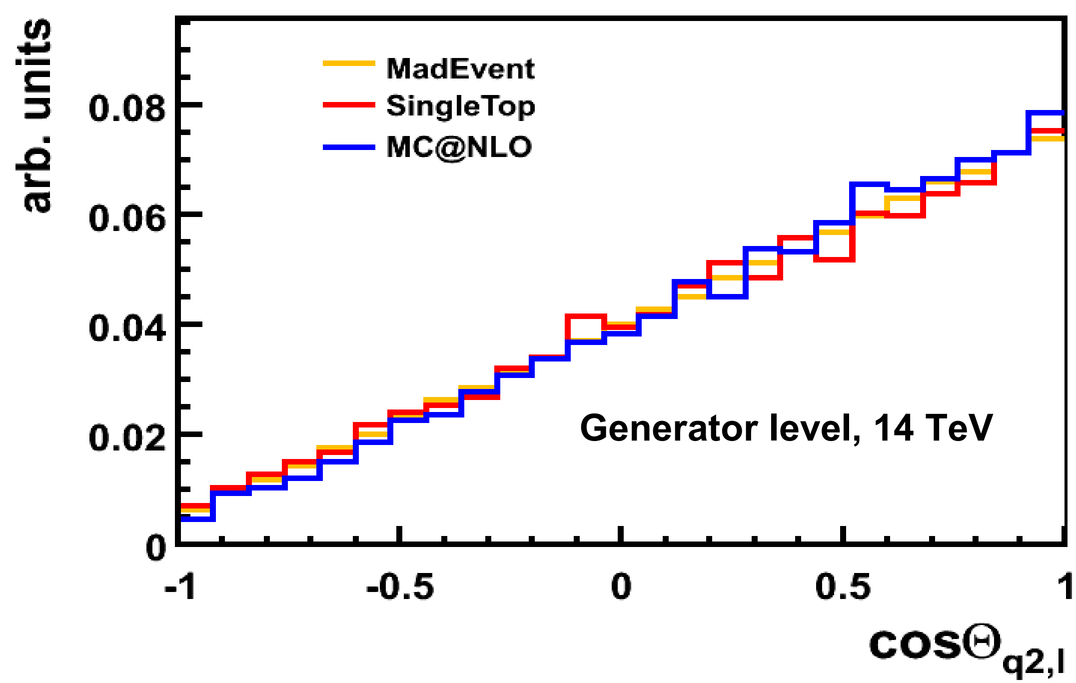
(top rest frame)



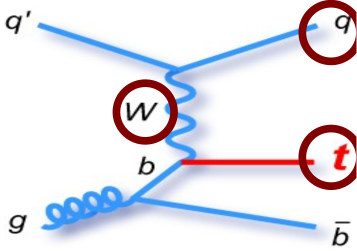
$$\frac{1}{\Gamma}\frac{d\Gamma}{d\cos\theta^*} = \frac{1}{2}(1 + A\cos\theta^*)$$

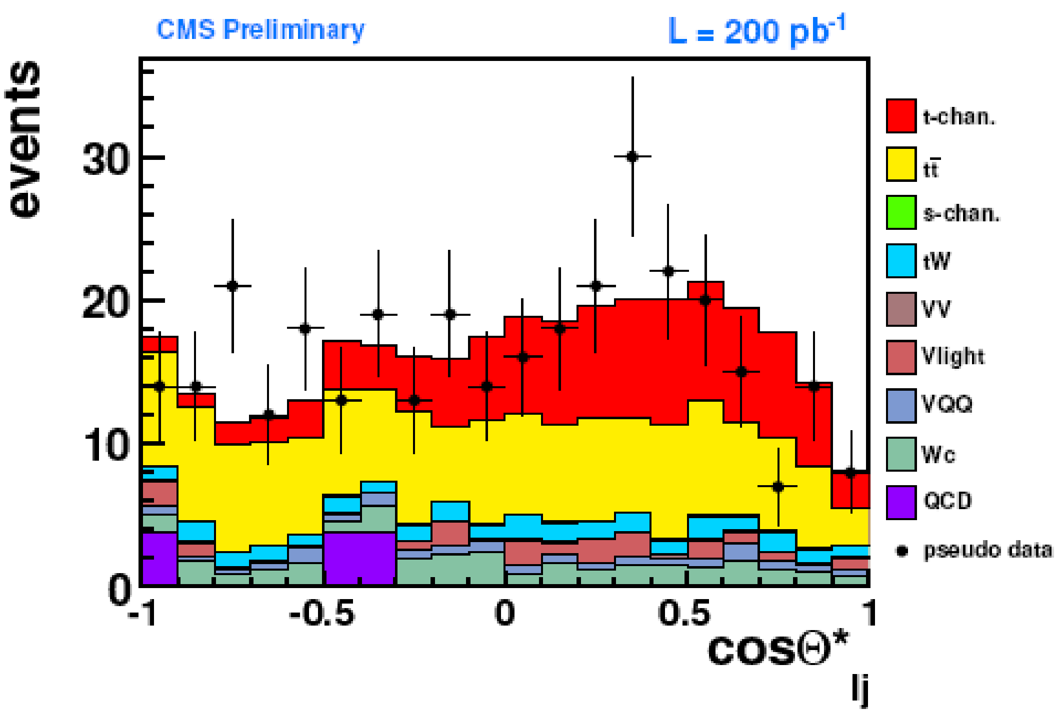
A=+1 for charged leptons

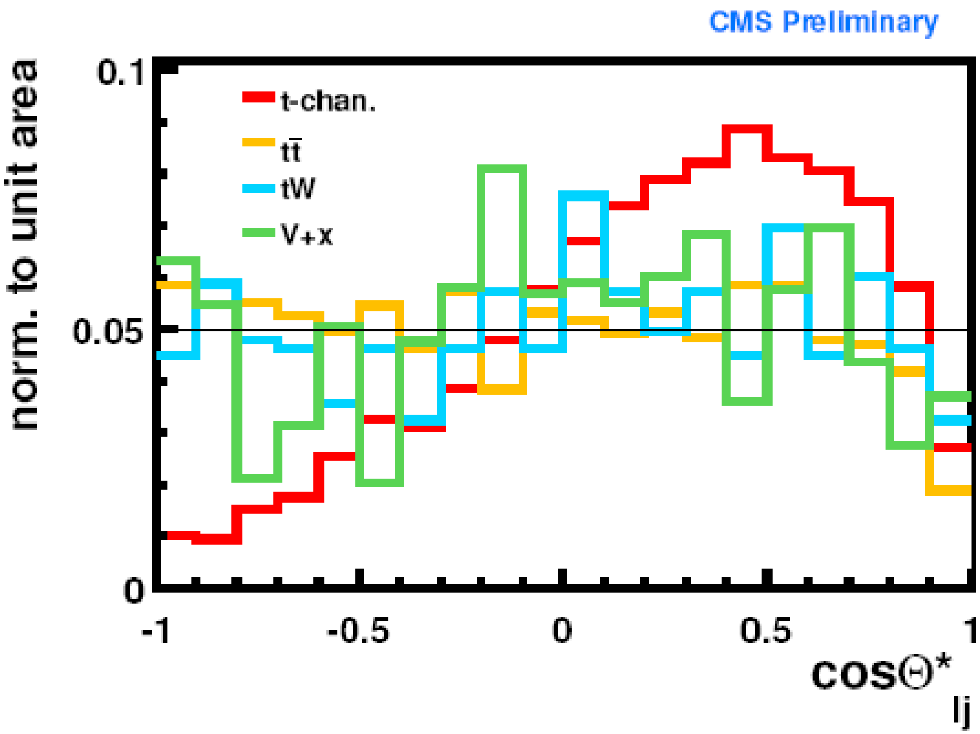




Smoking gun #4:
this top quark is polarized

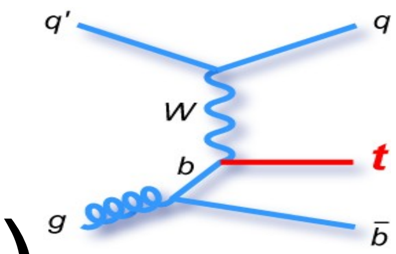






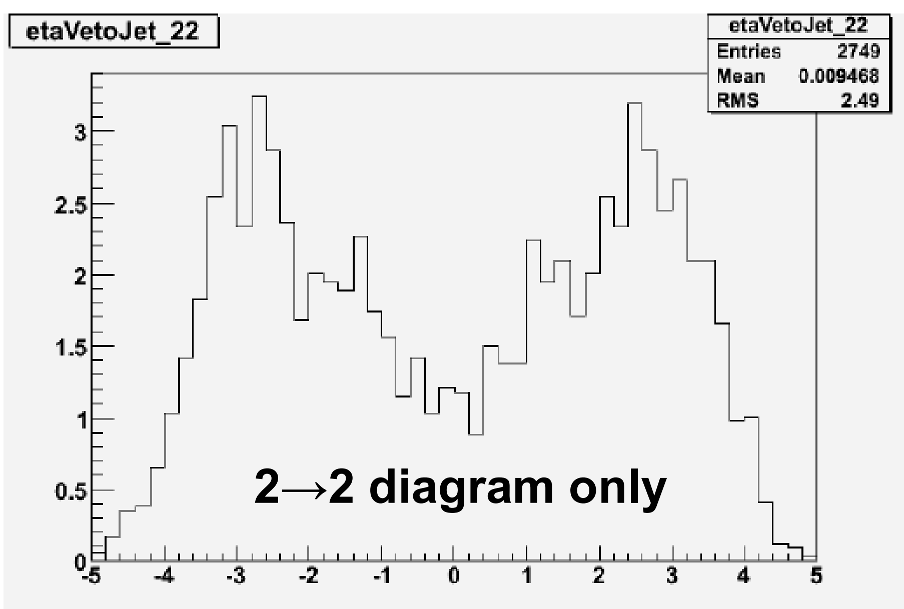
Backgrounds, instead, are remarkably flat...

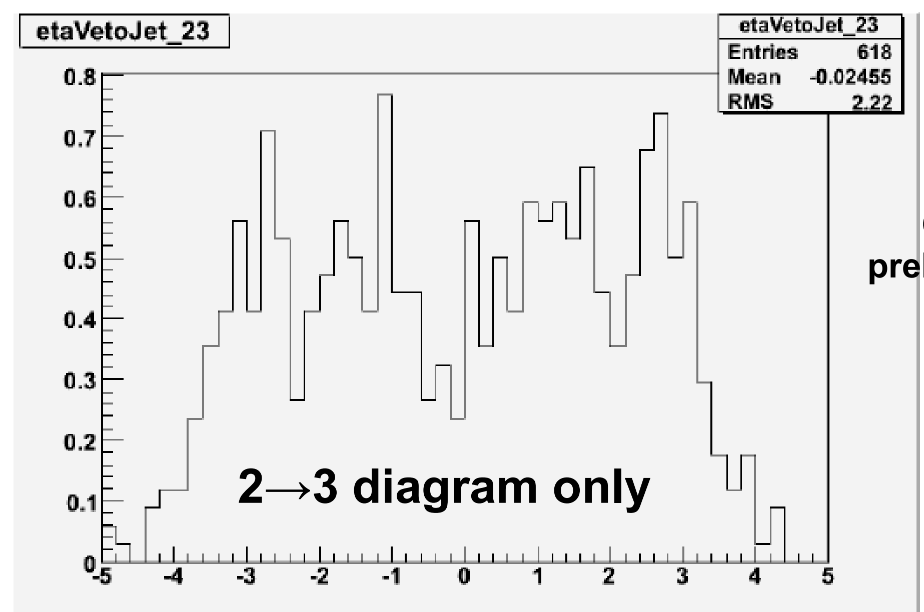
Un-flatness→signal



What to choose in a scenario of early data (1)

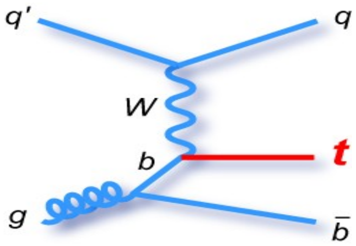
- Pseudorapidity of the recoil quark:
  - **Pro:** excellent discrimination against anything else; S/B>1 for |η|>2
  - **Contra:**
    - Sensitive to signal model
    - Relies critically on forward calorimetry; needs reliable understanding of forward jets, Underlying Event, Minimum Bias Events



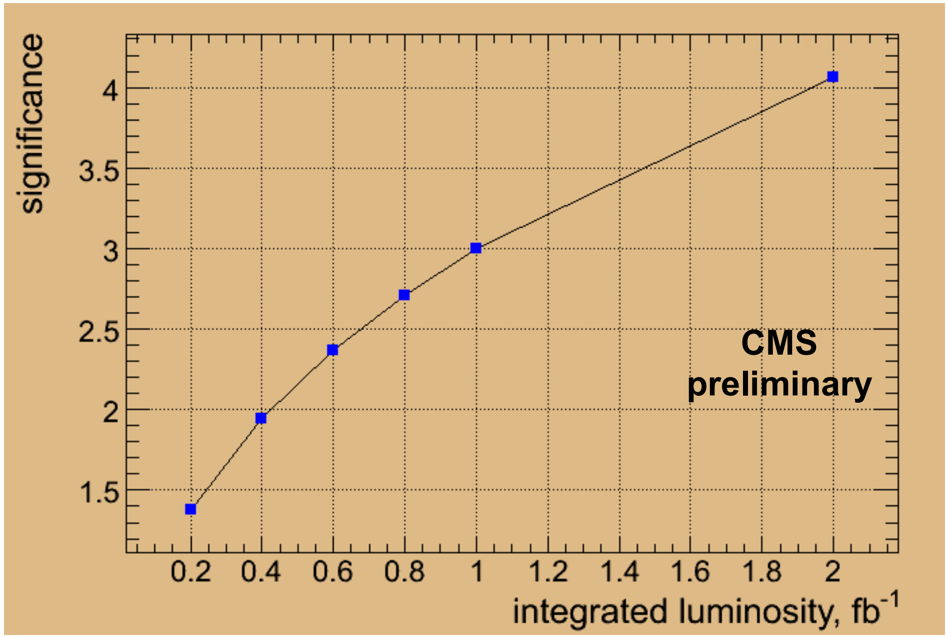


CMS preliminary

What to choose in a scenario of early data (2)

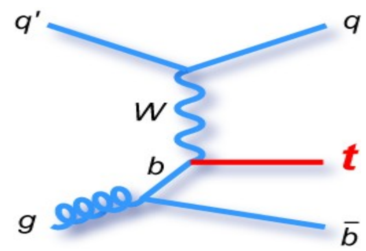


- Charge asymmetry:
  - **Pro:** most backgrounds and most systematics cancel away
  - **Contra:**
    - PDF systematic becomes critical
    - W+jets is charge-asymmetric too, thus it doesn't cancel out; simultaneous data-driven extraction of its **σ*A** is under consideration, but more work needed
    - most of all, statistical error is larger (N^+-N^- ~ N/4)

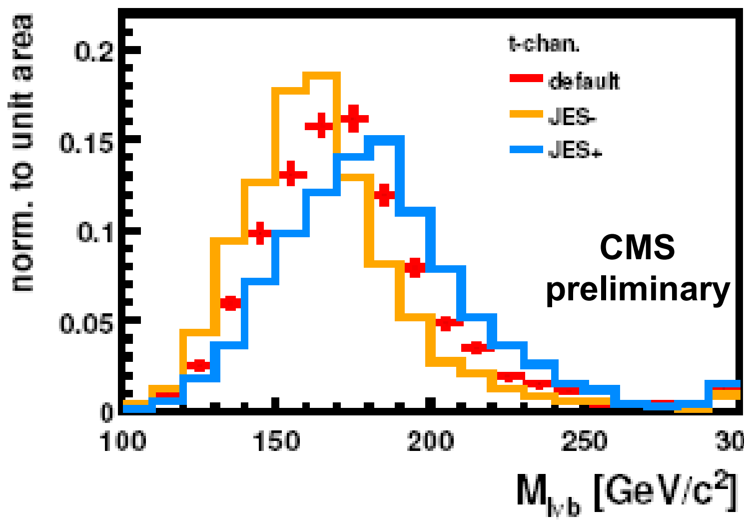


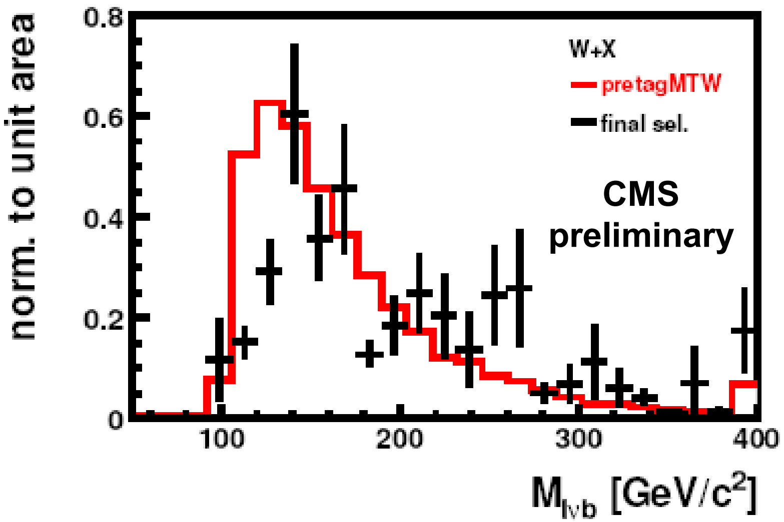
Muon channel only, same event selection, systematics included, uncertainty on the W asymmetry taken equal to what we expect after ~100/pb from the dedicated measurement in the 0j sample (and assumed 100% correlated for signal)

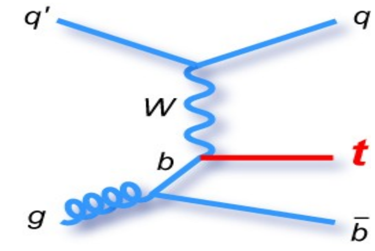
What to choose in a scenario of early data (3)



- Reconstructed mass:
  - **Pro:**
    - Boost-invariant
    - Good discrimination; we tried a template fit w/ 3 free parameters (st, tt & tW, W/Z+X; QCD constrained with the method seen before) and it works
  - **Contra:**
    - shape is very sensitive to jet uncertainties and gluon radiation
    - We tried to take W/Z+X and QCD shapes from control samples with relaxed selection, but corrections would be needed

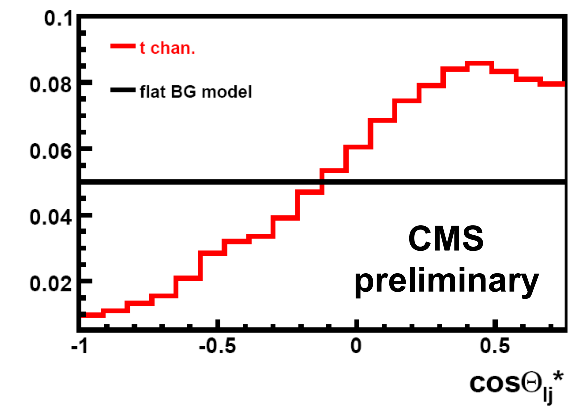


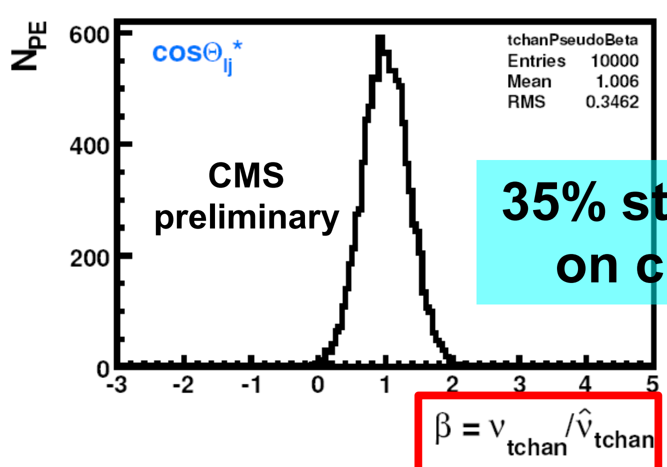




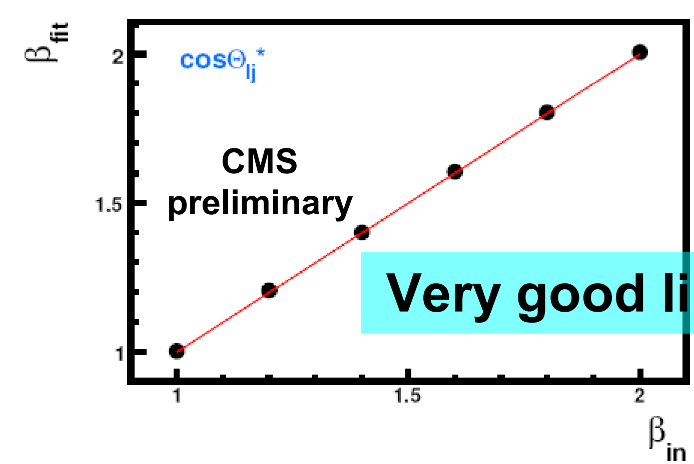
What we chose

- Binned likelihood fit based on **cosθ* in [-1,¾] range**
- Signal template taken from MC
- **Flat template assumed for sum of bkg**
- Free parameters: β_{signal}, β_{bkg} (β=measured/predicted)
- **No assumption about background size**
  - ±50% variation on bkg size → $\Delta\beta_{signal}$~0%, $\Delta\beta_{bkg}$~±50%



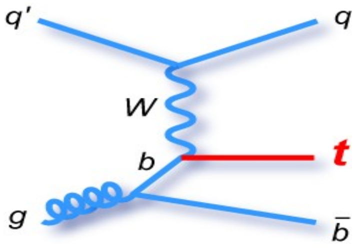


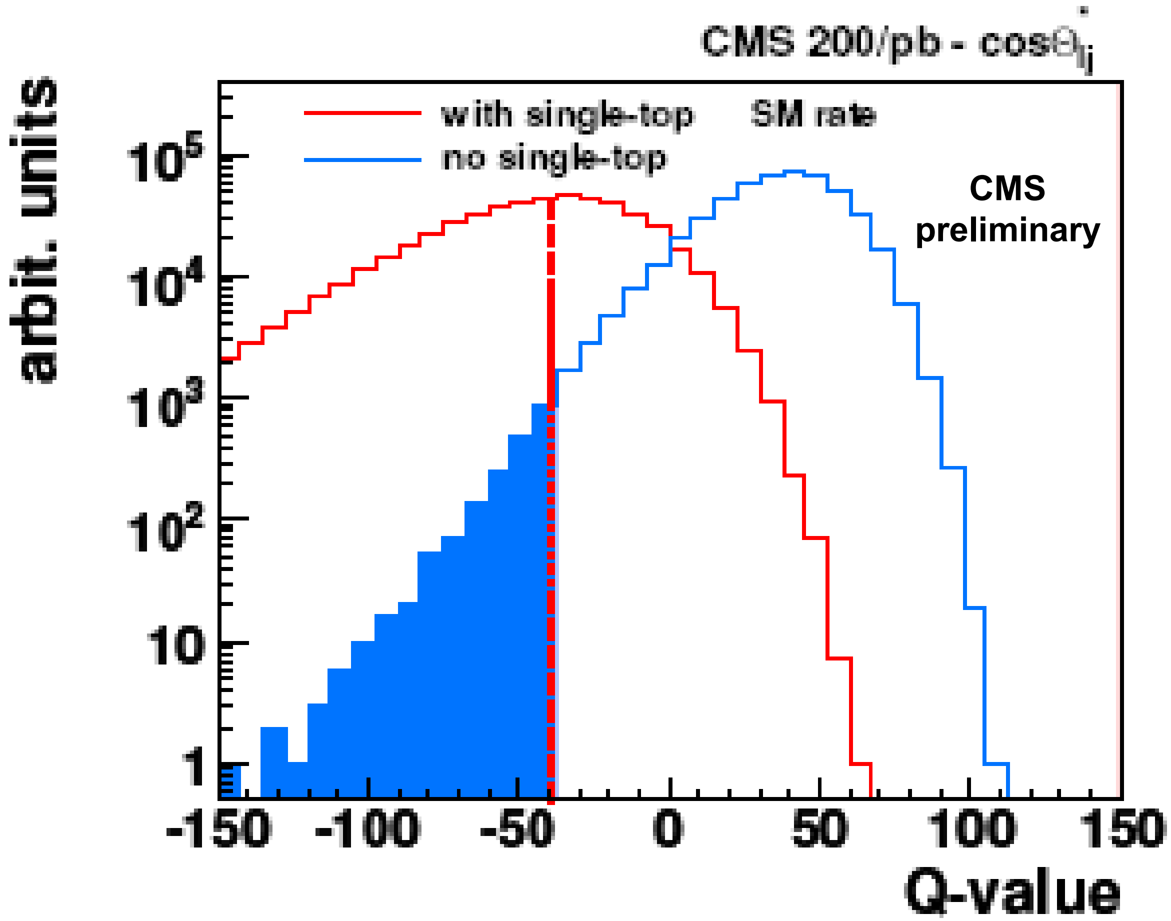
35% stat. uncertainty on cross section



Very good linearity

Expected sensitivity





$$Q = -2\ln\left(\frac{L(\hat{\beta}_1, \hat{\beta}_2, \ldots, \hat{\beta}_C)}{L(\beta_1 = 0, \tilde{\beta}_2, \ldots, \tilde{\beta}_C)}\right)$$

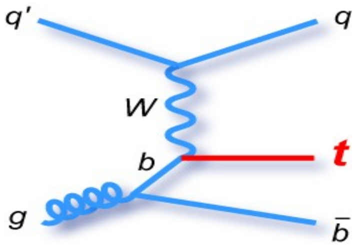
$$p(Q_m) = \frac{1}{A_q} \cdot \int_{-\infty}^{Q_m} q_0(Q')\, dQ'$$

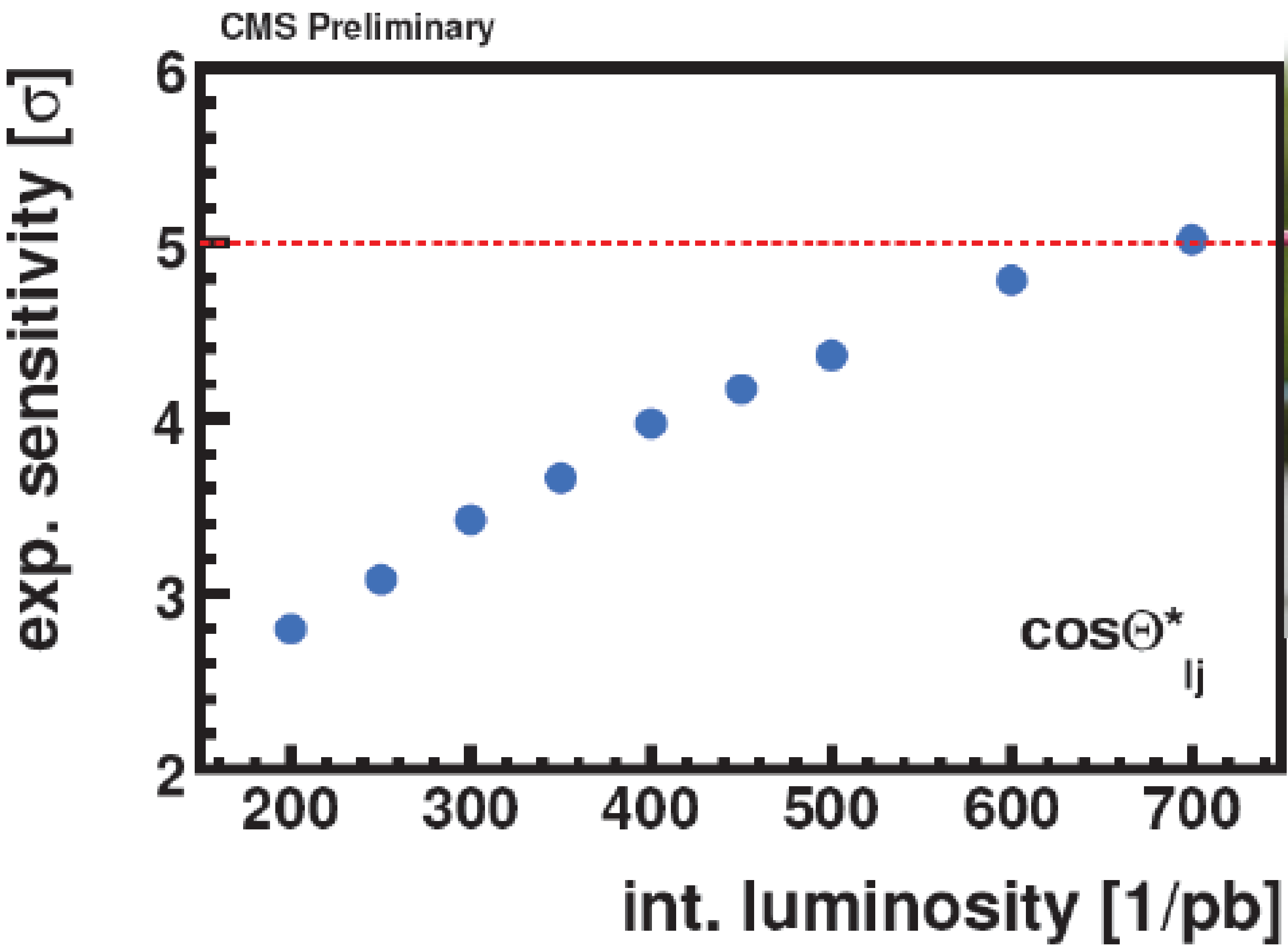
$$\sigma = \sqrt{2} \cdot Erf^{-1}(1 - 2(1 - p))$$

50k ensemble tests
("toy MC") performed

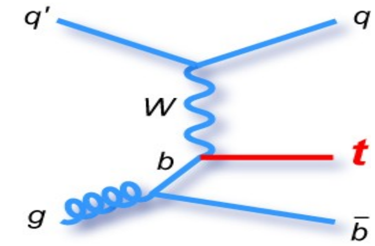
If both signal and background are described by the SM, there is a 50% probability of excluding the bkg-only hypothesis at 2.8σ level (stat.only) with 200/pb @ 10 TeV

And with more data?



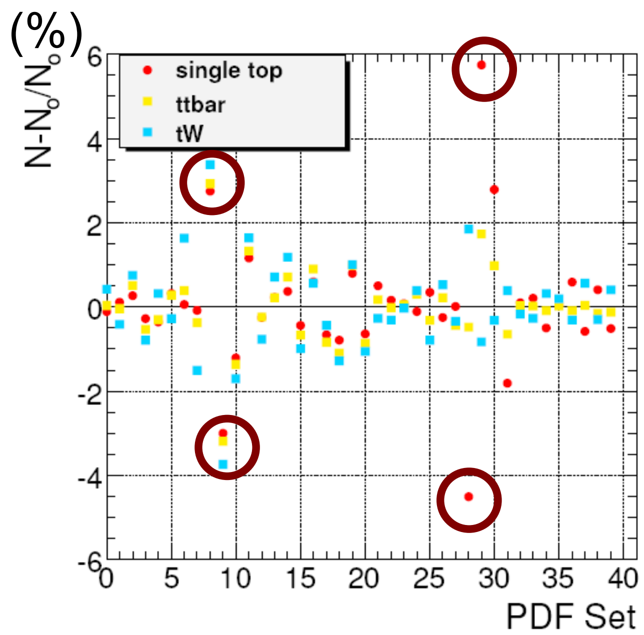


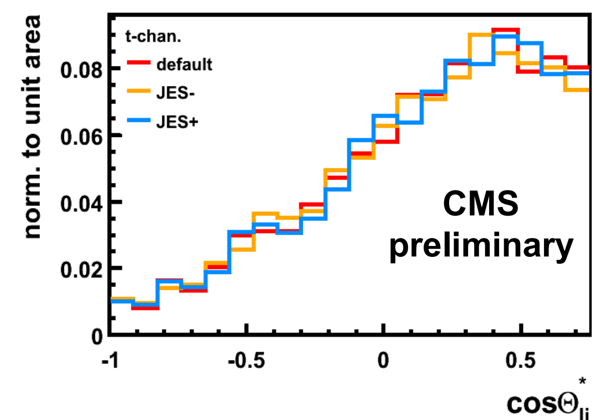
Systematics



- PDF (rate&shape):
  - CTEQ61 weights for signal and 2 major bkg's
  - Shape variations negligible in all cases
  - Rate variations up/down dominated by sets 28/29 (signal) and 8/9 (tt, tW) but always < 6%
- JES and MET (rate&shape):
  - ±10% on JES, propagated to the jet corr. in MET
  - ±10% on the uncorrected MET
- B tagging (rate&shape):
  - ±8.2%(8.0%) on efficiency of tight(loose) cut
  - ±18.1%(3.4%) on mistag prob. of tight(loose) cut
- Luminosity (rate): ±10%

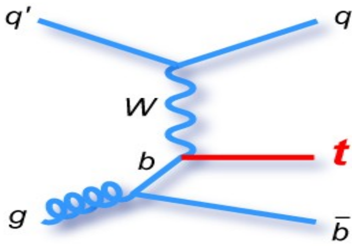
Considering all systematics: 2.8σ → 2.7σ





Plenty of shape variation plots in the backup slides

Cross section precision and its limiting factors

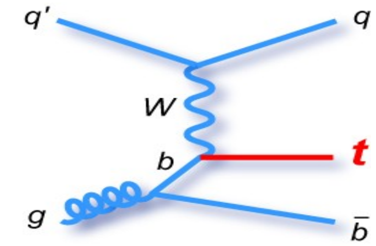


Source of uncertainty	$\Delta\sigma$ [%]	Expected sensitivity
statistical	± 35	2.8σ
b tagging	± 7.3	2.7σ
mistag	± 0.4	2.7σ
JES	± 5.5	2.7σ
MET	± 9.9	2.7σ
PDF	± 5.5	2.7σ
total	± 39	2.7σ

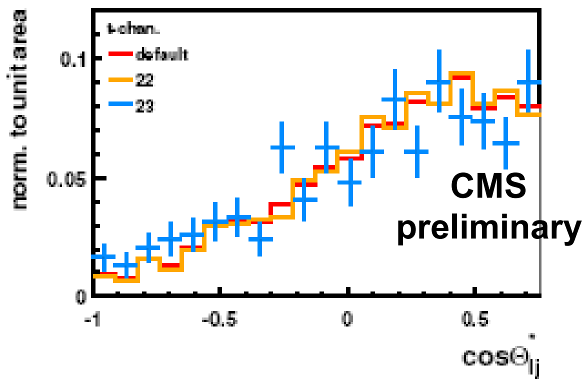
(±10% luminosity)

- At 200/pb, by far limited by statistics
- By the time of 1/fb, data-driven methods are expected to improve the knowledge of these sources of systematics as follows:
  - JES uncertainty from ±10% to ±5%, MET *probably* the same
  - b-tagging uncertainty from ±8% to ±5-6%
  - PDF uncertainties reduced by large factors, see e.g. http://cdsweb.cern.ch/record/1117860/files/ATL-SLIDE-2008-079.pdf
  - luminosity uncertainty from ±10% to ±5%

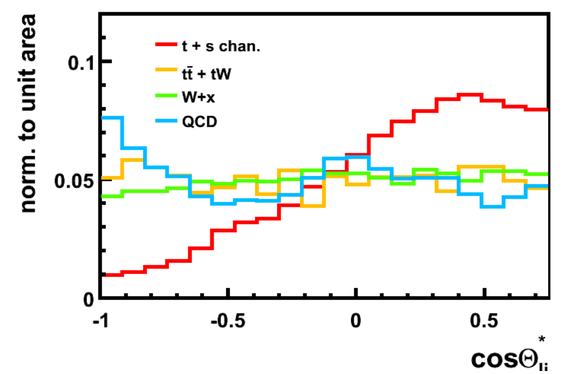
“What happens if...”



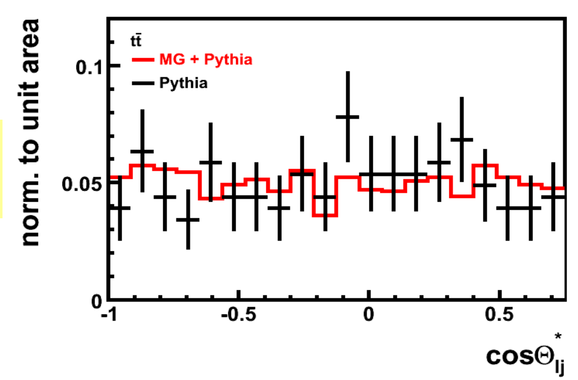
- Signal model (shape): the most extreme variation, a priori, is 2→2(only) vs 2→3(only)



- Bkg model (shape): the observed shape is used instead of the flat assumption
  - tt+tW: shape taken from MC
  - W/Z+X: shape from b-tag-less control sample
  - QCD: from b-tag-less <u>anti-isolated</u> control sample

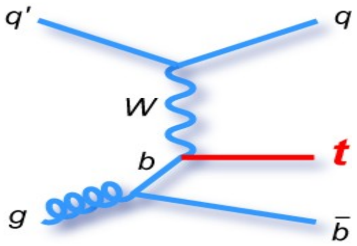


- Radiation model for ttbar (shape):
  - ISR/FSR up and down
  - MadGraph vs Pythia



The worst difference (QCD shape): 2.7σ → 2.6σ

- Overall background rate +/-50%: no significant bias on the measurement, sensitivity 2.2σ / 3.2σ

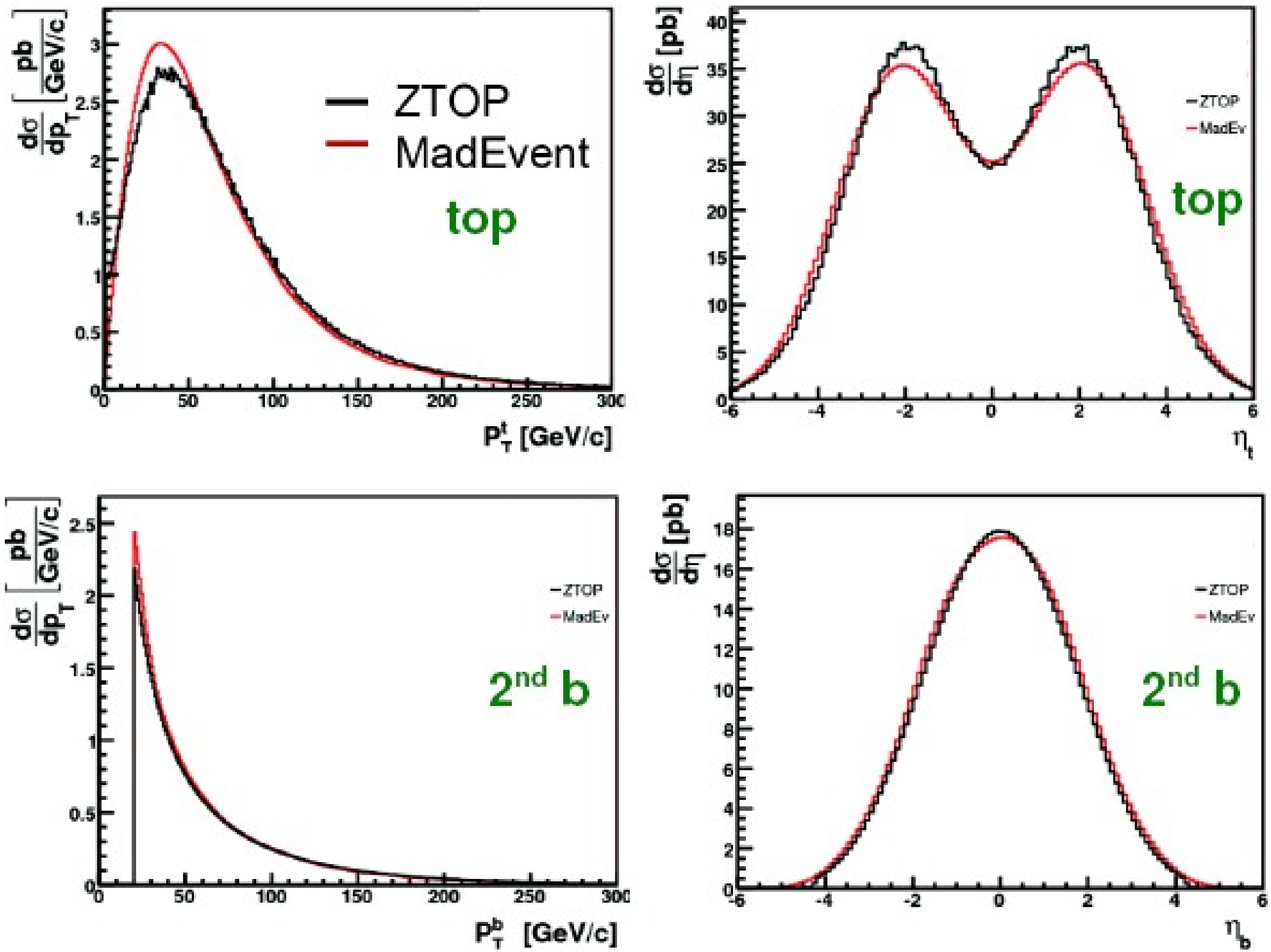


Conclusions

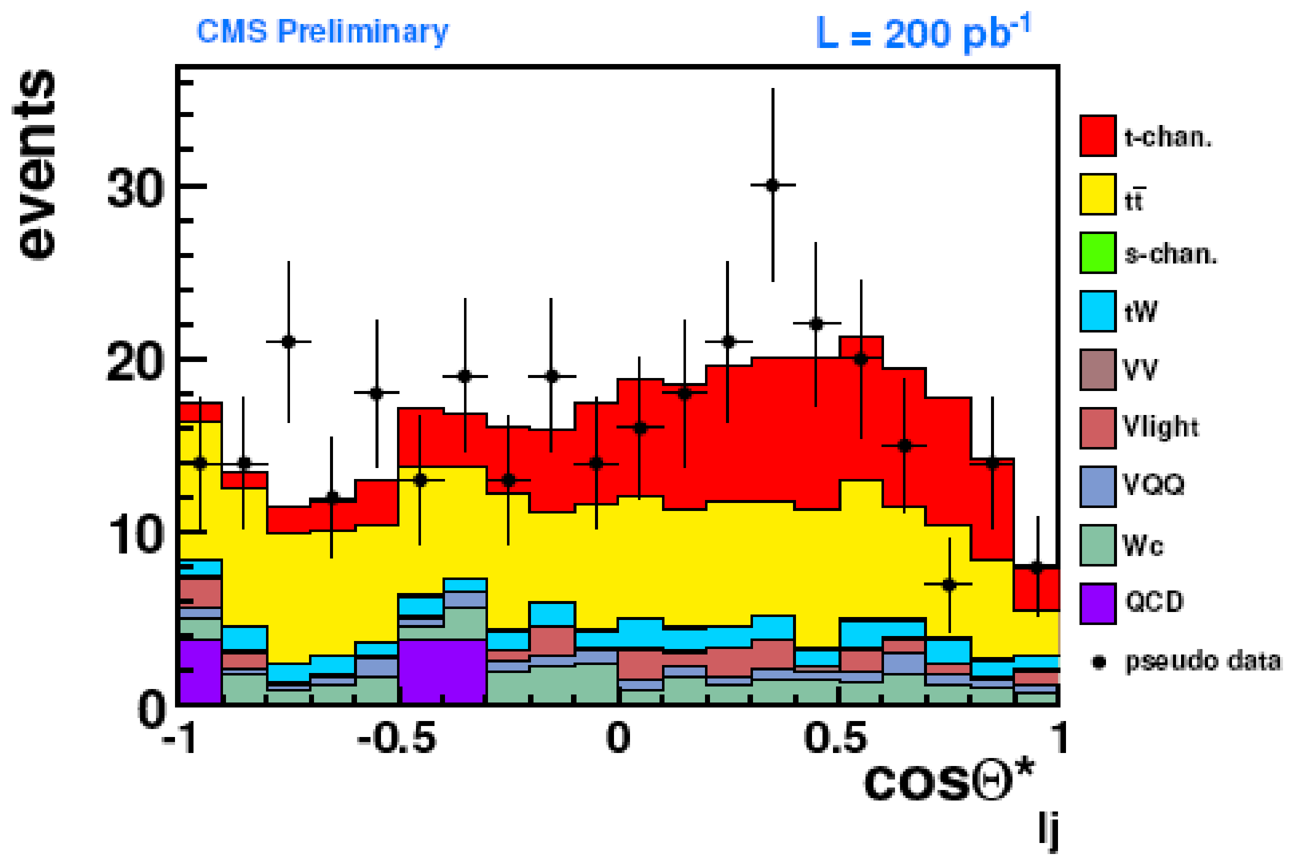
- **New method** based on fitting the muon-jet angle in the reconstructed top rest frame
  - Flatness of the **overall background**: no a priori assumption on its size, treated as a **free parameter**
  - Assumption of flatness for individual backgrounds can be monitored with control samples
  - Make sure that there are no surprises by QCD bkg: *in situ* estimation from the full M_T distribution
- This method is **robust against systematics**
- Plans for first long physics run are for **~200/pb @~10 TeV:** can realistically achieve **~3σ**

Backup slides

Signal model: 2→2 / 2→3 matching



Hunting for new physics: first constrain the backgrounds!



This variable is sensitive to FCNC and anomalous Wtb couplings.
Ideally, independent precise measurements of all SM backgrounds would permit to measure the non-SM component of single top from the remaining pedestal.
But this use is not for early data.

Scaling to 7 TeV

- Cross sections, w/ MCFM (NLO, m_t=170 GeV, CTEQ6M):
  - Single top, t: 85.4+47.3 pb (10 TeV) → 42.3+21.9 pb (7 TeV)
  - Wt: 27.3 pb (10 TeV) → 11.1 pb (7 TeV)
  - Pair production: 468.6 pb (10 TeV) → 186.7 pb (7 TeV)
  - Wc: 3.3 **n**b (10 TeV) → 1.9 **n**b (7 TeV)
  - Wbb: 29.9+19.1 pb (10 TeV) → 16.8+10.1 pb (7 TeV)
  - Naïve rescaling of $S/\sqrt{B}$ for the cosθ* method: 200/pb @ 10 TeV → ~450/pb @ 7 TeV; I ignored QCD (12 ev @10 TeV), W+light jets (12 ev @10 TeV), and all minor bkg's (9 ev@10 TeV)

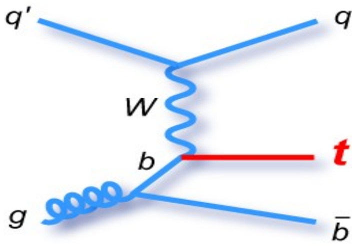
Planned analysis improvements

- Combine cosθ^* and charge asymmetry
- Add electron channel
- Particle flow
- Kinematic fit

Why M_T instead of MET

- Better discrimination power
- Better stability vs JES and MET variations
- Easier QCD estimation: all non-QCD processes have a similar M_T shape, not so for MET
- In QCD, MET is correlated with muon momentum and muon isolation (M_T is not), due to the fact that most of the surviving QCD events are bb or cc
  - Probably not true for the electronic channel

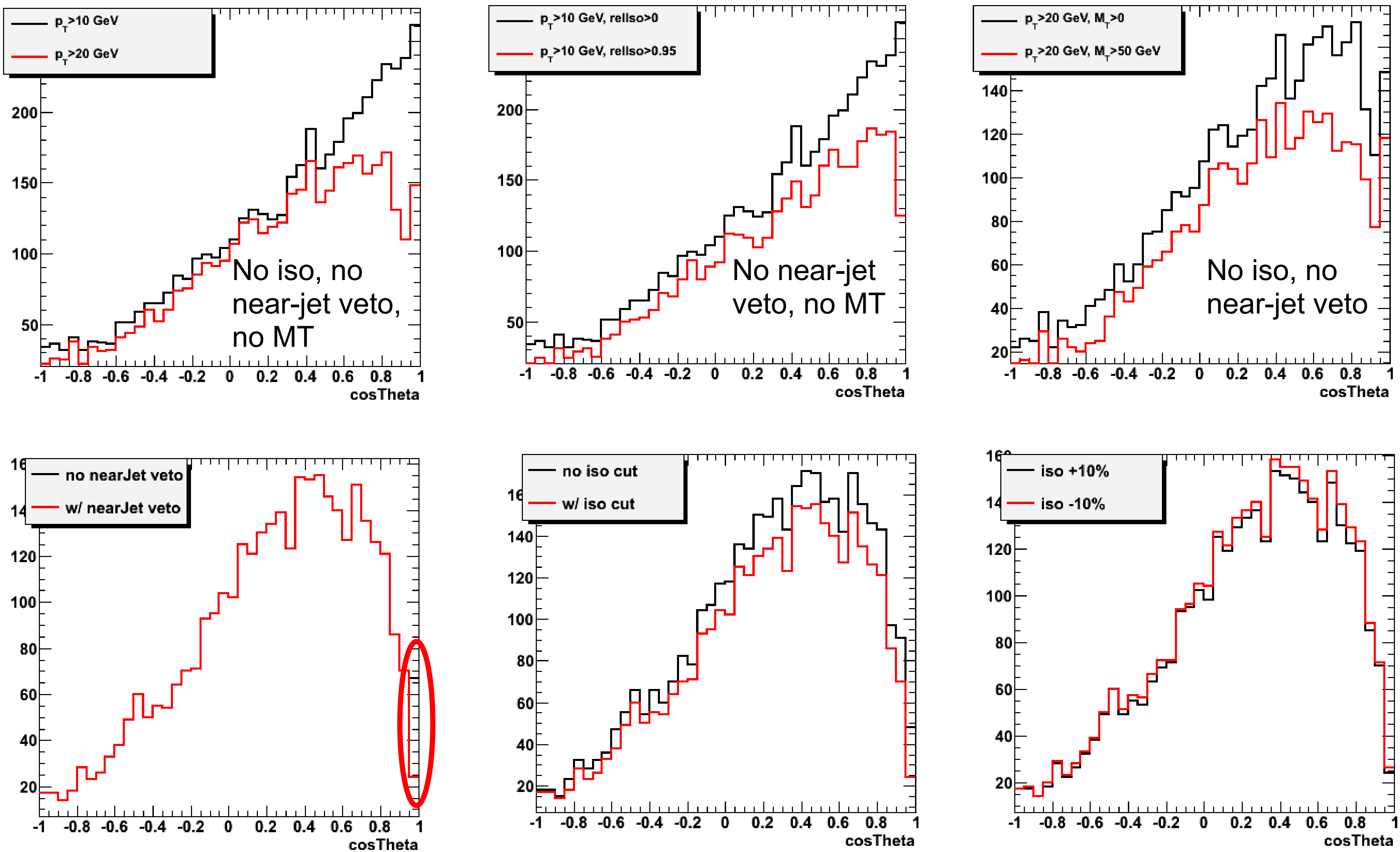
What to choose in a scenario of early data (4)



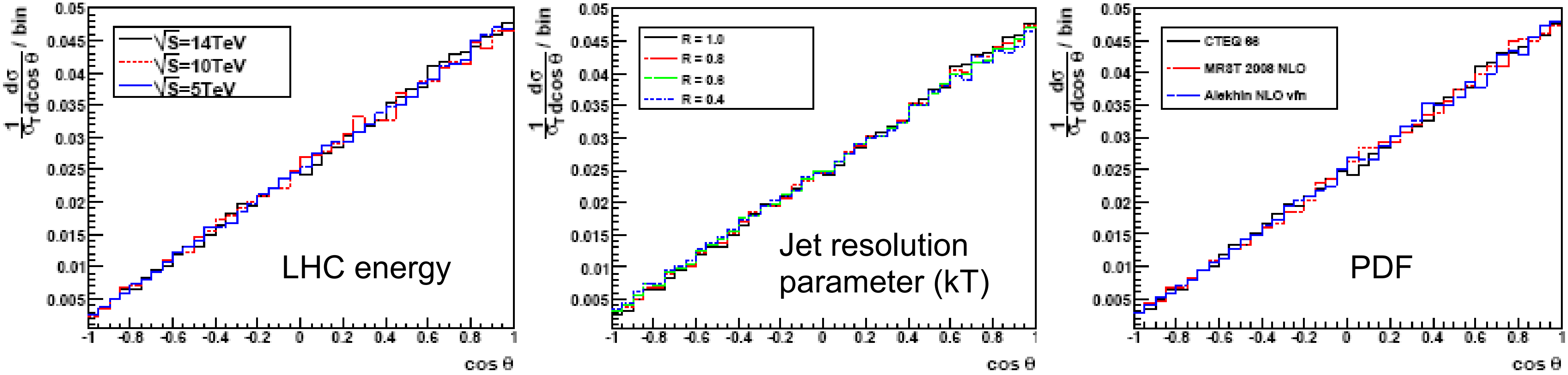
- Polarization:
  - **Pro:**
    - All backgrounds share the same shape (and it is a very simple one!)
    - Shape is remarkably stable against theory and detector systematics, for both signal and backgrounds
  - **Contra:**
    - Close to ~1, sensitive to kinematic cuts and isolation
    - Complication when used in conjunction with η cut: bias on bkg makes it more signal-like

What shapes cosθ*

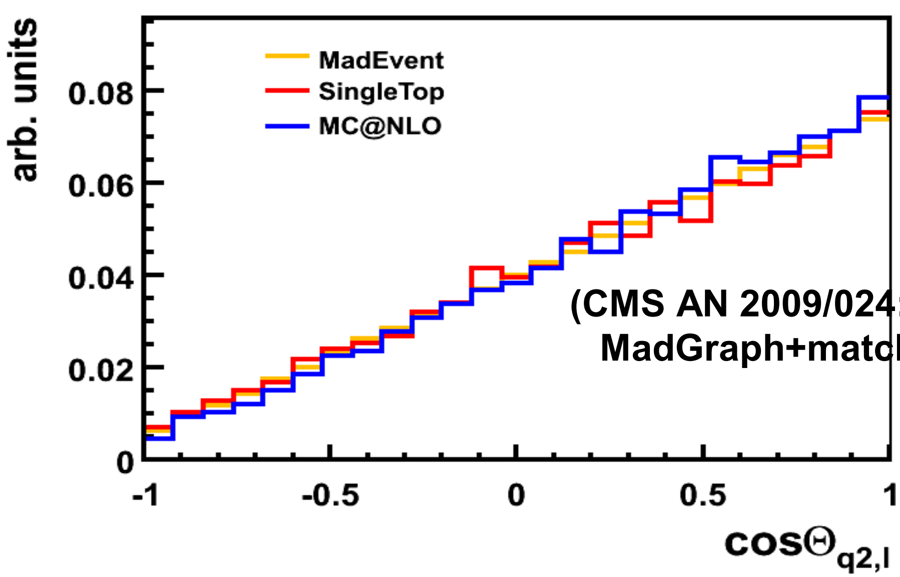
CMS preliminary



At generator level



(P.Motylinski, hep-ph:0905.4754)



(CMS AN 2009/024: 14 TeV, comparison between MadGraph+matching / SingleTop / MC@NLO)

W charge asymmetry
(CMS-PAS-EWK-09-003)

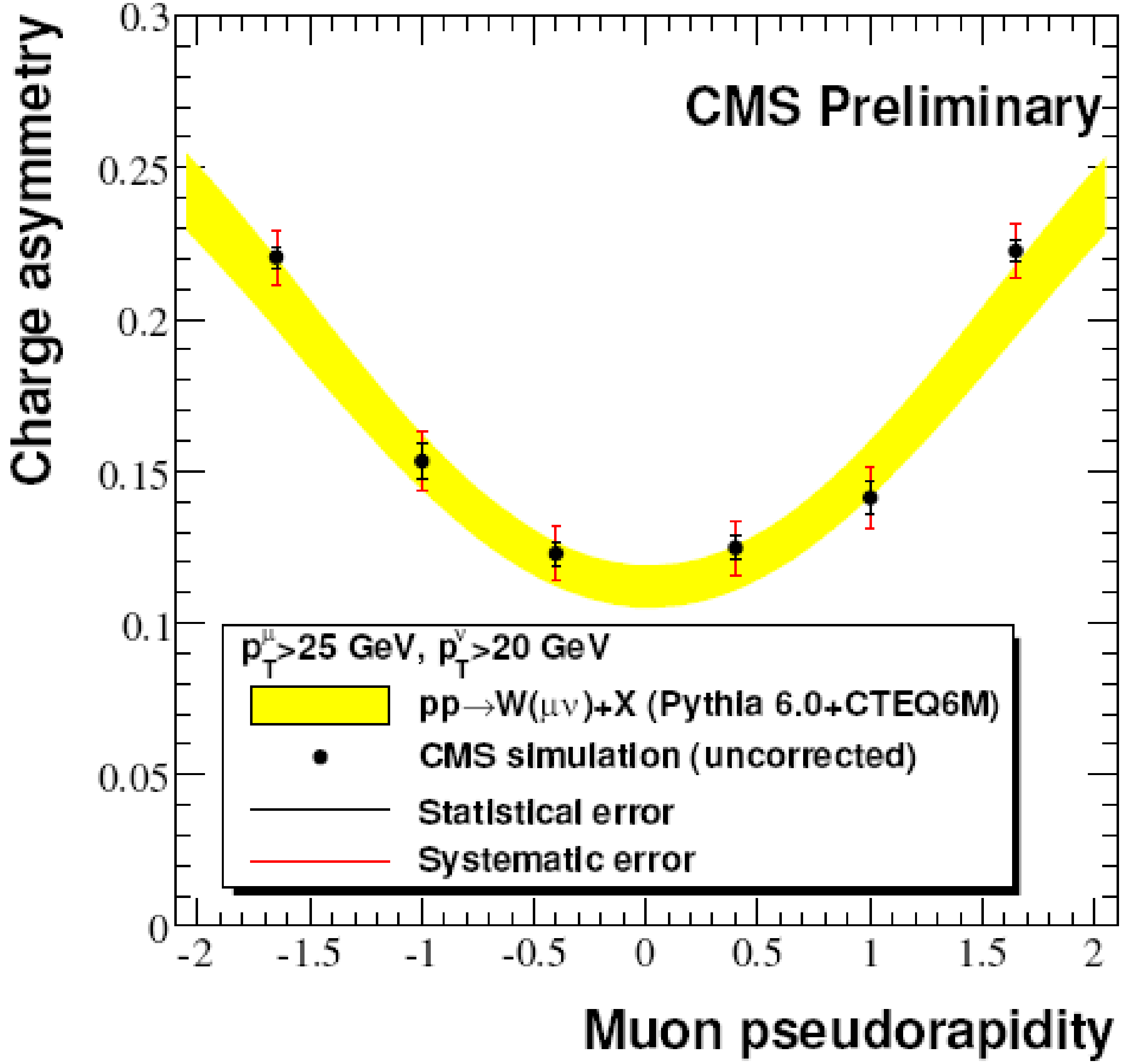
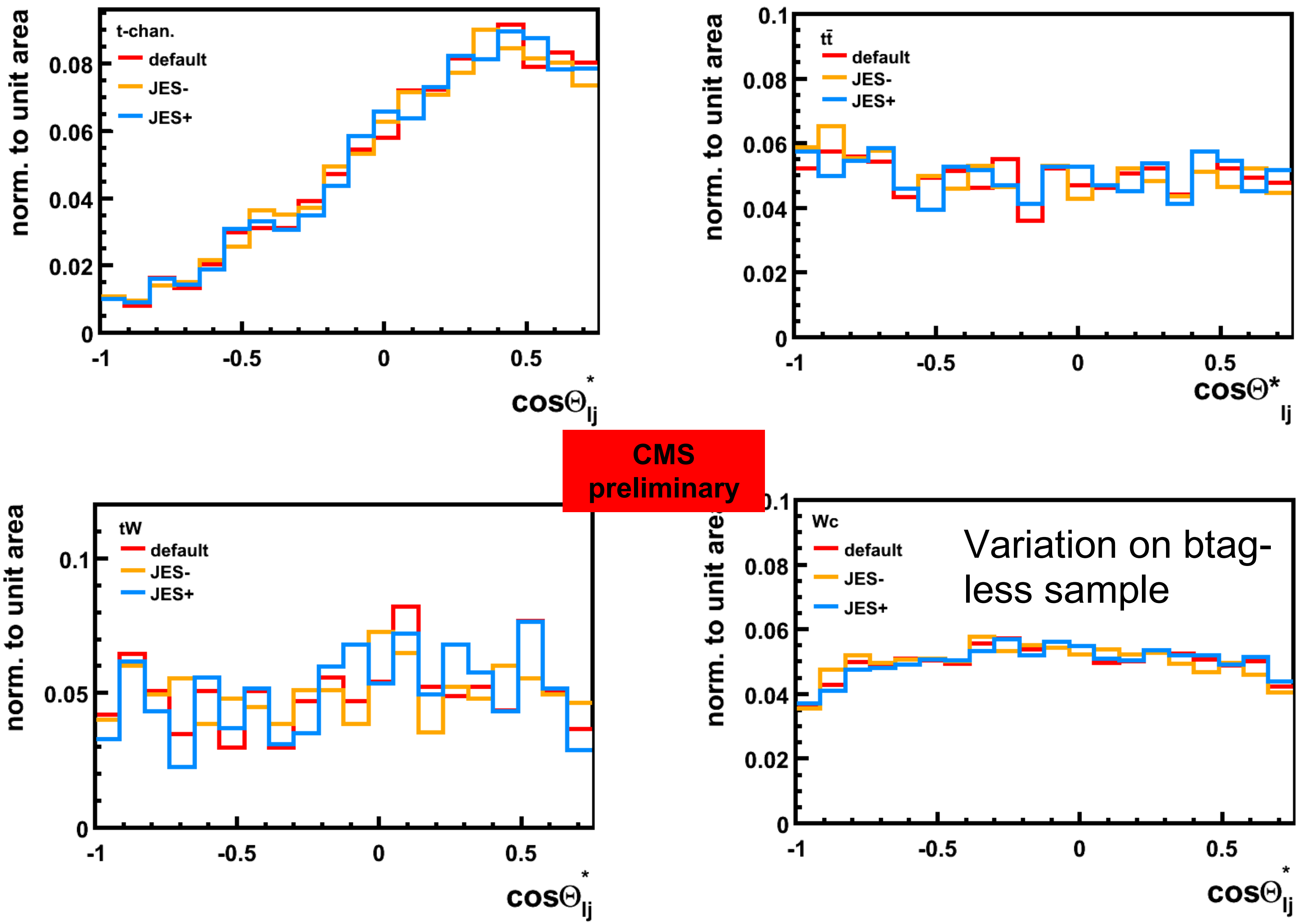
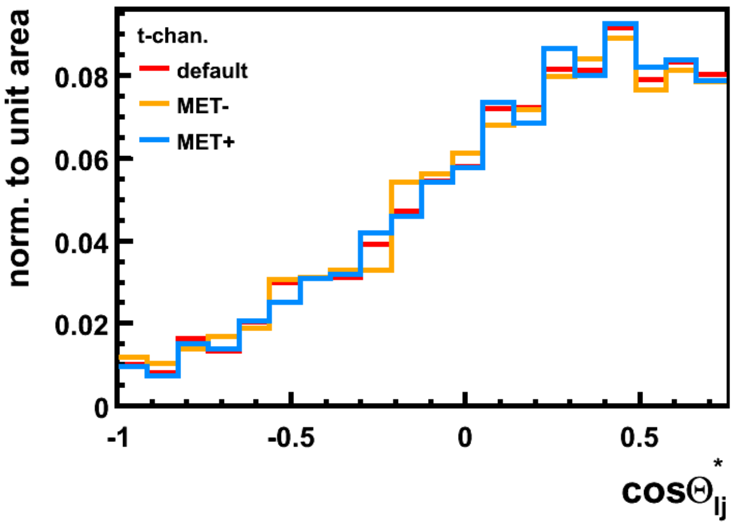


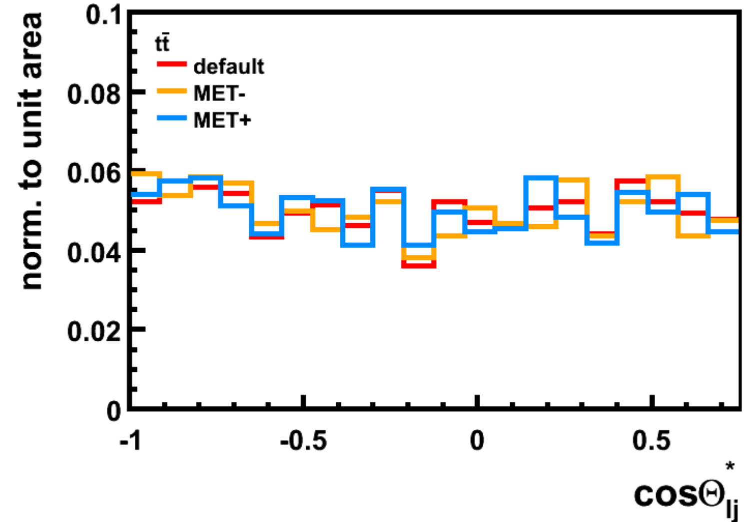
Figure 5: The measured charge asymmetry for 100 pb^{-1} of simulated luminosity. The systematic error is dominated by the statistical error on the efficiency ratio determined using 100 pb^{-1} of Drell-Yan MC.

Shape systematics, JES

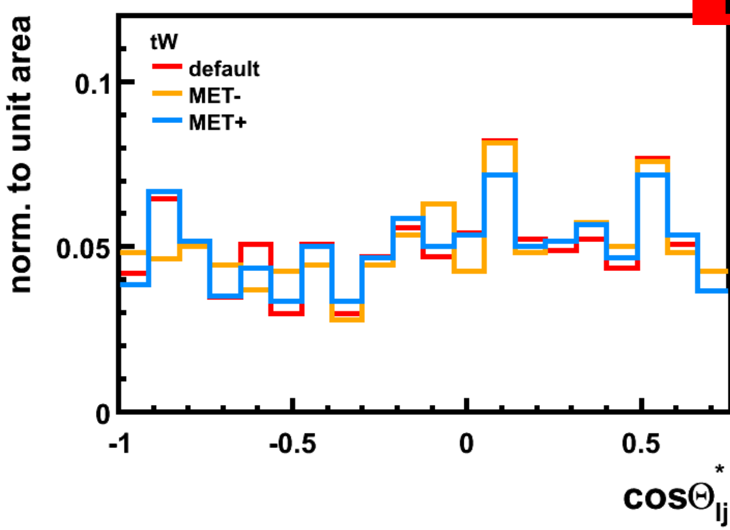


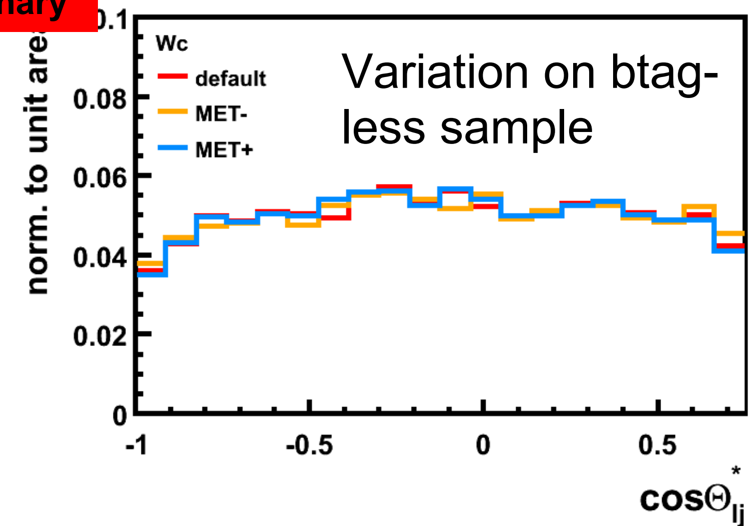
Shape systematics, MET



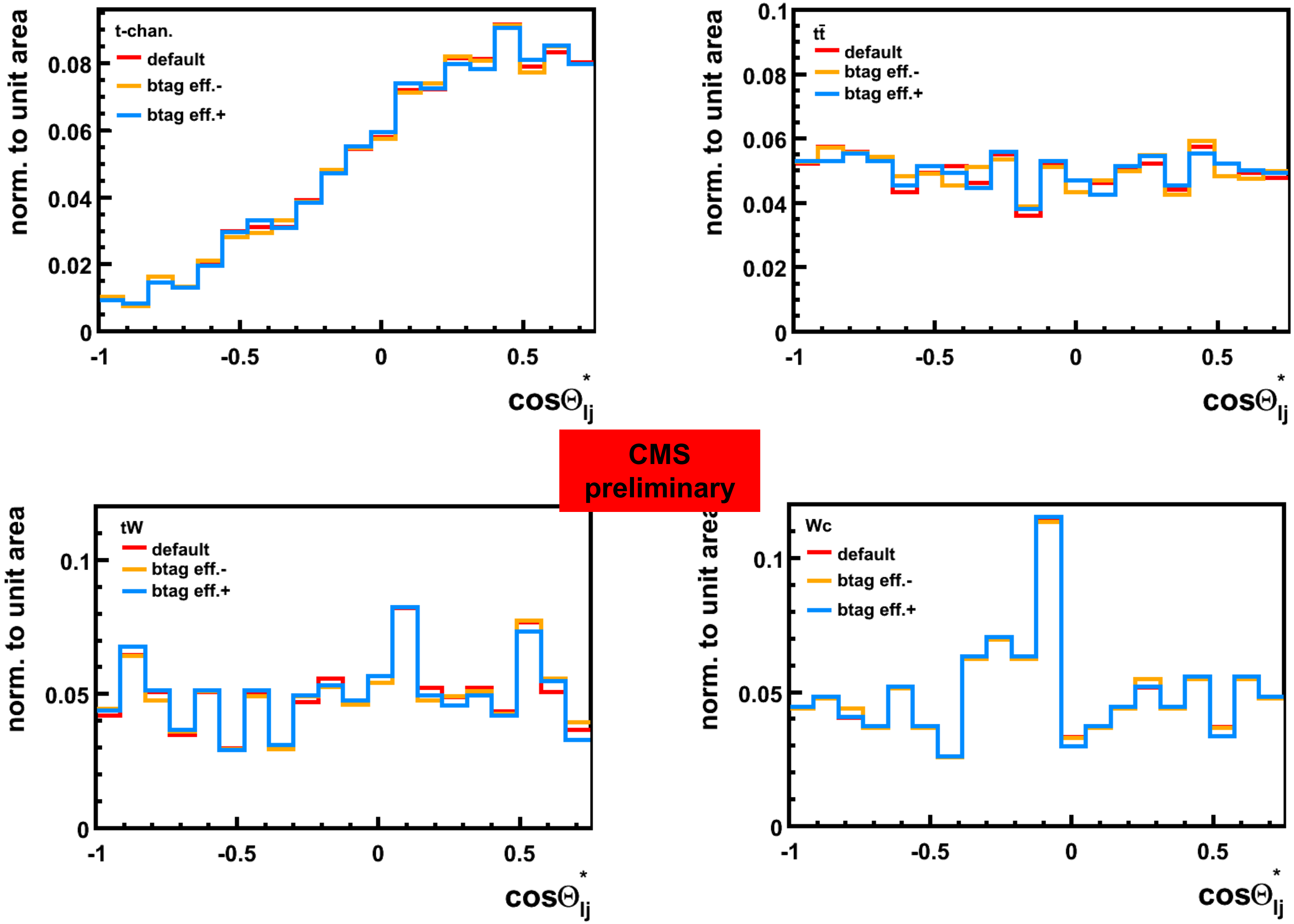


CMS preliminary

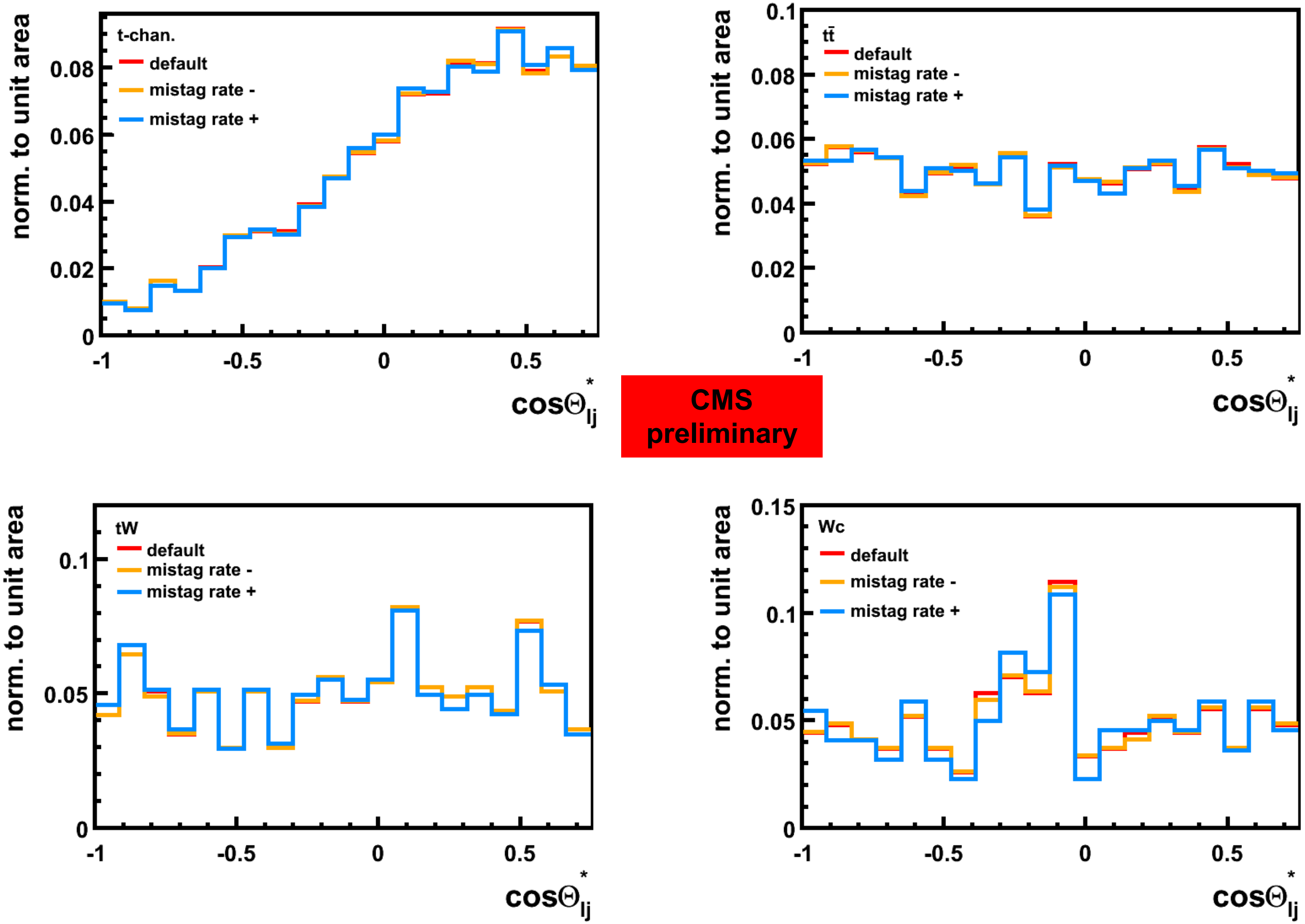




Shape systematics, b tagging



Shape systematics, mistag



Btag efficiency

- From CMS-PAS-BTV-07-001:

operating point	Loose			Medium			Tight		
Luminosity (pb^{-1})	10	100	1000	10	100	1000	10	100	1000
Systematics (%)									
β	5.8	5.8	2.9	6.3	6.3	3.2	5.7	5.7	2.9
α	0.4	0.4	0.2	0.4	0.4	0.2	0.4	0.4	0.2
κ_b	3.4	3.4	1.7	3.6	3.6	1.8	3.3	3.3	1.7
κ_{cl}	0.2	0.2	0.1	0.2	0.2	0.1	0.2	0.2	0.1
p_{Trel}	2.8	2.8	2.8	2.9	2.9	2.9	3.0	3.0	3.0
statistics MC (%)	2.3	2.3	2.3	2.6	2.6	2.6	2.7	2.7	2.7
statistics data (%)	7.2	2.3	0.7	8.4	2.6	0.8	8.7	2.7	0.9
Total error (%)	10.5	8.0	6.4	11.8	8.6	5.4	11.6	8.2	5.3

Table 6: Summary of uncertainties expected for b-tagging efficiencies measured with the System8 for different luminosity scenarios for the TrackCounting tagger.

Mistag efficiency

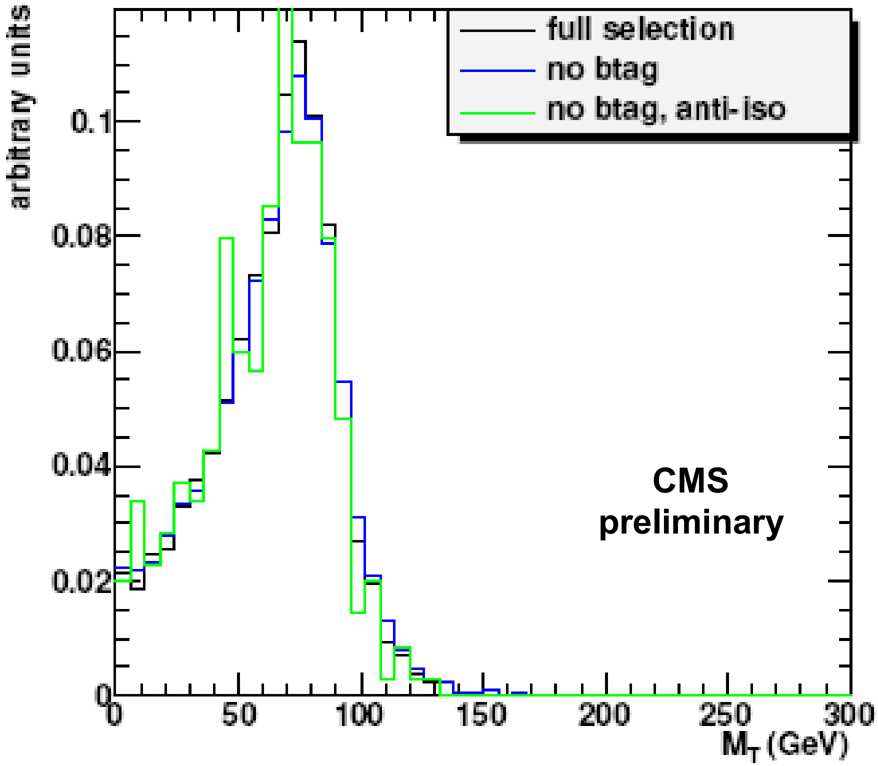
- From CMS-PAS-BTV-07-002:

The breakdown of systematic uncertainties is detailed in Table 2. As, according to Figures 8-13 and 24, the systematics depend on the jet p_T and η, they are reported here for $p_T = 100$ GeV, integrating over all η.

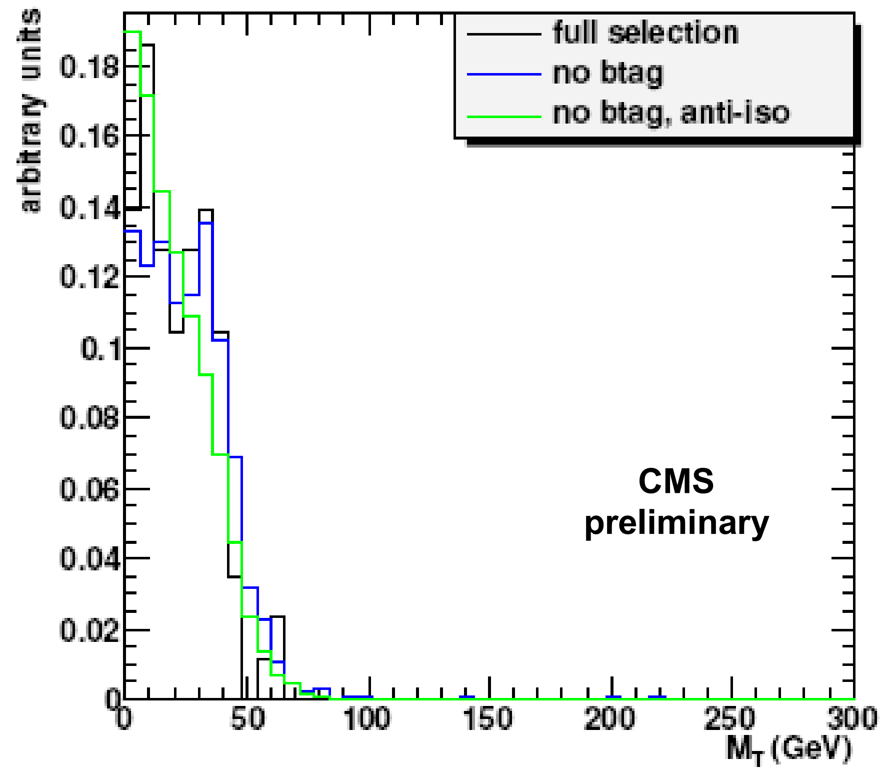
operating point	Loose			Medium			Tight		
Luminosity (pb^{-1})	10	100	1000	10	100	1000	10	100	1000
Systematics (%)									
b fraction	1.4	1.4	0.6	0.8	0.8	0.3	1.2	1.2	0.5
c fraction	0.8	0.8	0.3	0.7	0.7	0.3	1.3	1.3	0.5
g fraction	0.8	0.8	0.4	1.4	1.4	0.7	2.3	2.3	1.2
V^0 fraction	1.4	1.4	0.7	3.6	3.6	1.8	4.6	4.6	2.3
other displaced processes	1.4	1.4	0.7	3.6	3.6	1.8	4.6	4.6	2.3
IP sign flip	0.7	0.3	0.2	4.5	1.9	1.4	24.0	10.2	7.6
statistics MC	0.1	0.1	0.1	0.4	0.4	0.4	1.2	1.2	1.2
statistics data	0.4	0.1	—	1.6	0.5	0.2	5.5	1.7	0.6
sampling	2.0	2.0	2.0	5.0	5.0	5.0	13.0	13.0	13.0
Total syst.	3.4	3.4	2.4	8.8	7.6	5.9	28.7	18.1	15.5

Table 2: Estimated relative systematics (%) on the mistag efficiency for the Track Counting tagger at a jet $p_T = 100$ GeV. Three operating points, corresponding to an average mistag efficiency (in the QCD 80-120 Monte Carlo) of 10%, 1%, 0.1%, respectively, and three luminosity scenarios are considered.

QCD estimation: (in)sensitivity of M_T to btag & iso



Single top, t channel



Multi-jet QCD

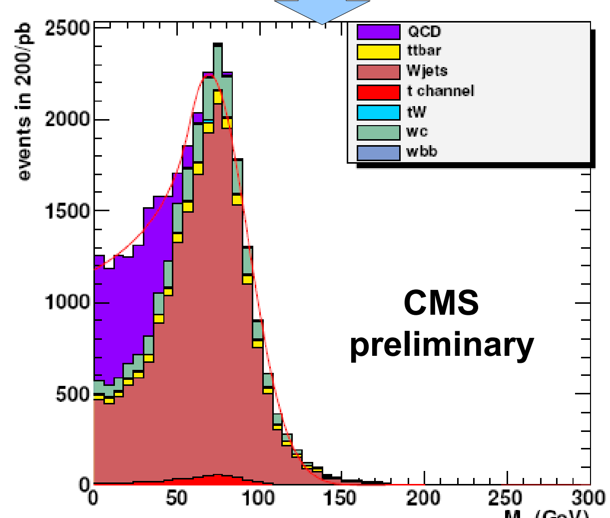
QCD-, W-, Z-enriched control samples: event yields

QCD-enriched

Process	N_{evt} in MC	N_{evt} in 200 pb^{-1}
QCD	56,920	222,036
signal	352	10
$t\bar{t}$	384	30
tW	118	4
$W+$ light partons	417	340
Wc	282	28
$Wb\bar{b}$	21	1

W-enriched

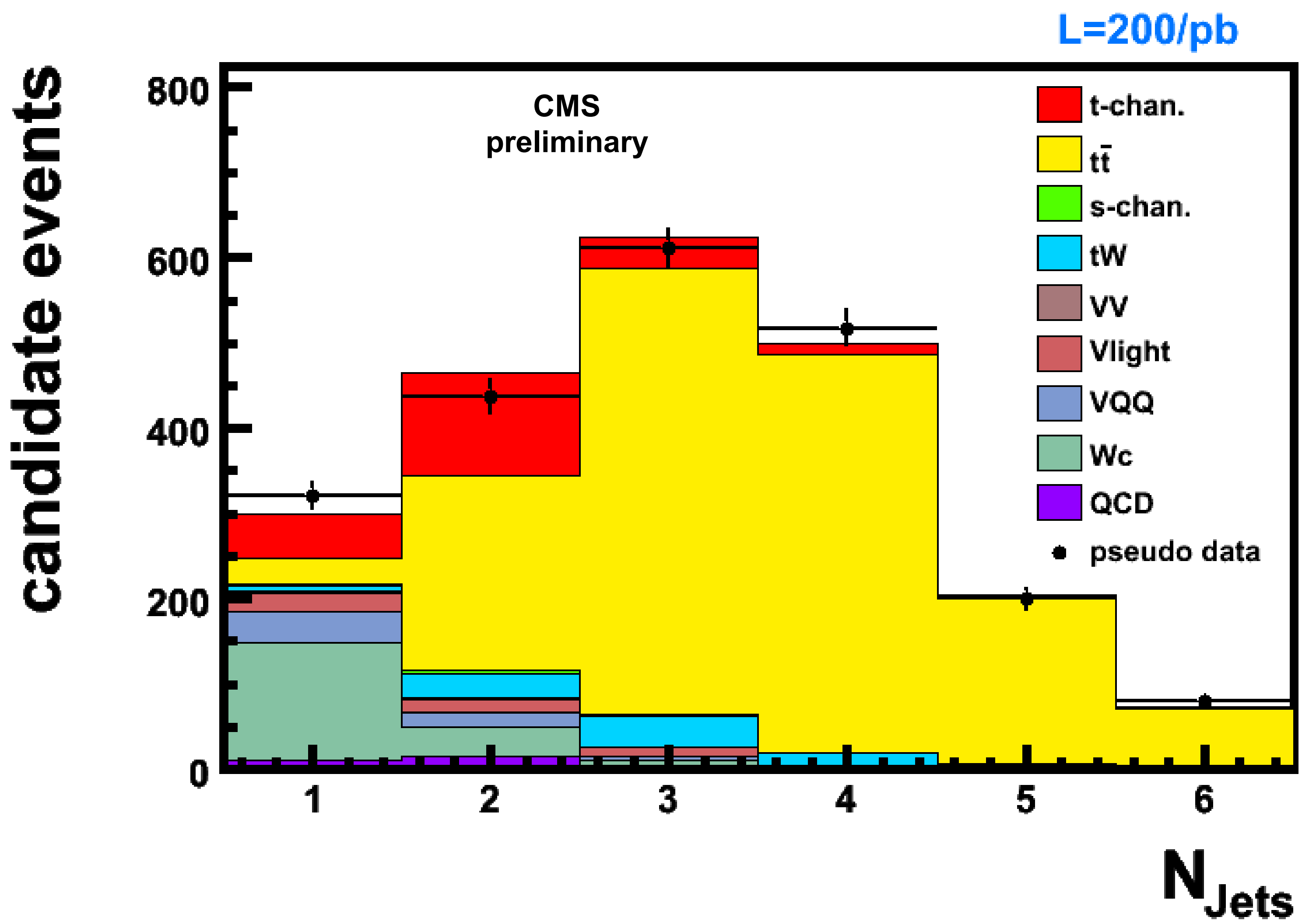
Process	N_{evt} in MC	N_{evt} in 200 pb^{-1}
QCD	1,342	5,235
signal	18,240	544
$t\bar{t}$	10,528	845
tW	4,379	150
$W+$ light partons	23,815	19,439
Wc	25,941	2,567
$Wb\bar{b}$	1,165	80



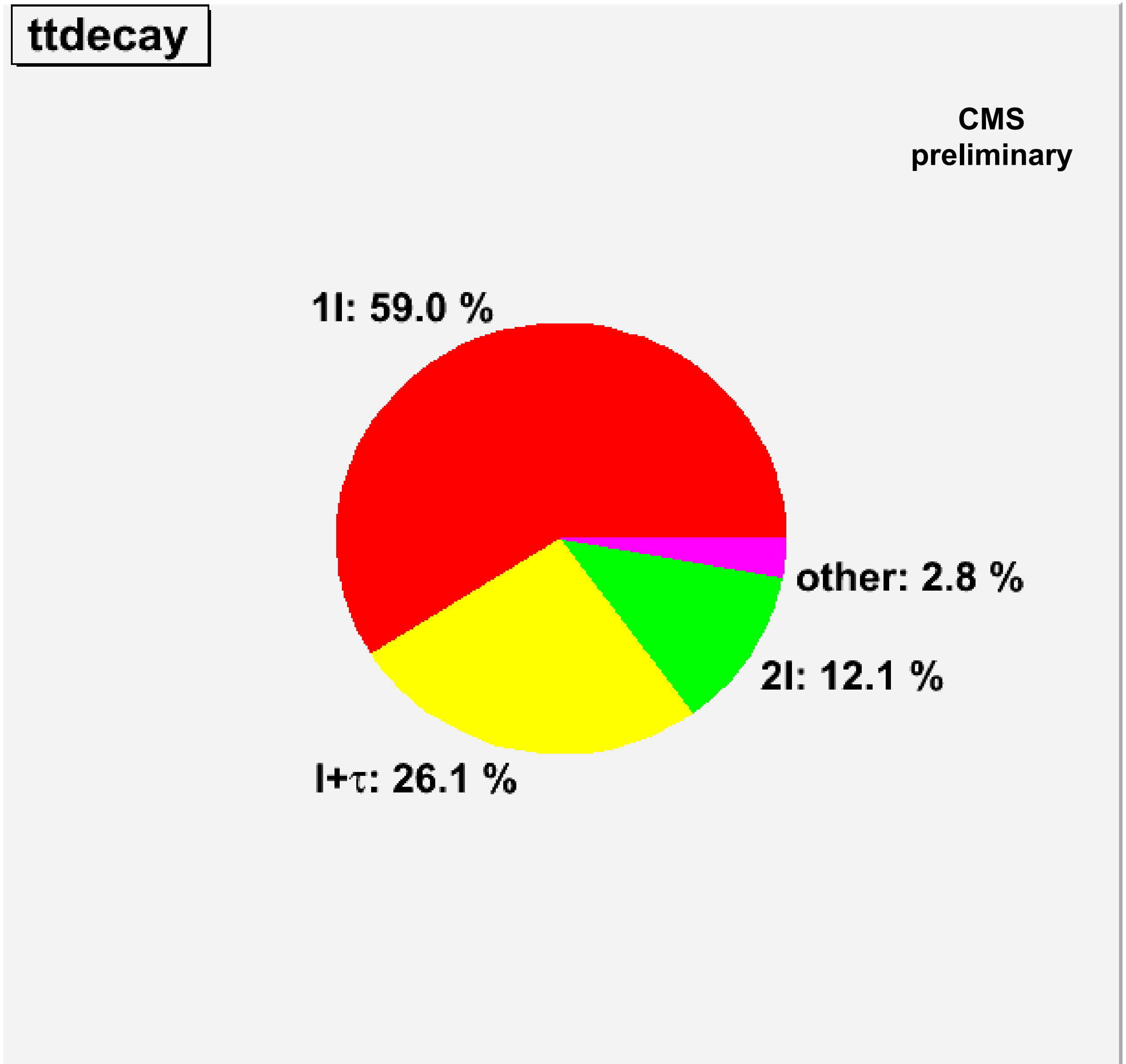
Z-enriched

Process	N_{evt} in MC	N_{evt} in 200 pb^{-1}
$Z+jets$	2,732	1,677
$Zc\bar{c}$	1,198	48
$Zb\bar{b}$	791	45
QCD	1	4
signal	76	2
$t\bar{t}$	530	48
$W+$ light partons	4	3

Full selection apart from 2nd b veto



Know your enemy: what kind of ttbar remains



Latest public studies by ATLAS
(CERN-OPEN-2008-020)

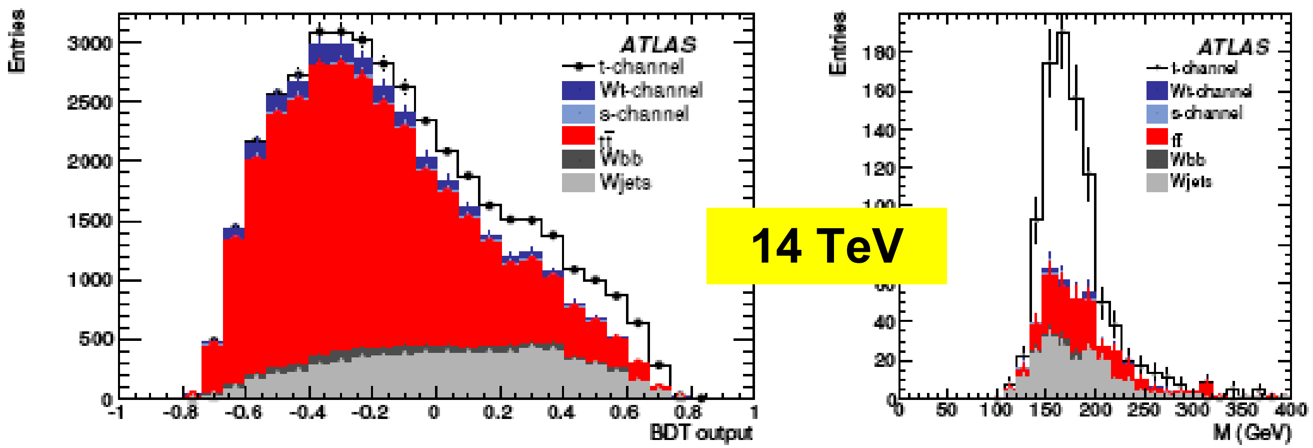


Figure 3: Boosted decision tree output for signal and background after the b-tagged jet p_T cut (left) and leptonic top quark mass distribution using cut on BDT output at 0.6 (right).

Source	Analysis of 1 fb^{-1} Variation	Cut-based	BDT
Data Statistics		5.0%	5.7 %
MC Statistics		6.5 %	7.9%
Luminosity	5%	18.3 %	8.8%
b-tagging	5%	18.1 %	6.6%
JES	5%	21.6%	9.9%
Lepton ID	0.4%	1.5 %	0.7%
Trigger	1.0%	1.7 %	1.7%
Bkg x-section		22.9%	8.2%
ISR/FSR	+7.2 -10.6%	9.8 %	9.4%
PDF	+1.38 -1.07 %	12.3 %	3.2%
MC Model	4.2%	4.2 %	4.2%
Total		45%	22%

Several topological variables as input to the BDT; cosθ* is not used

Results are consistent with the old (TDR2006) CMS analysis @ 14 TeV when the same scenario is considered for statistics and systematics

Analyses in a 200/pb @ 10 TeV scenario are currently in progress

Digression: *a 4th family? Wasn't it excluded since long time?*

Electroweak precision data [LEPEWWG]

- Particle Data Group:

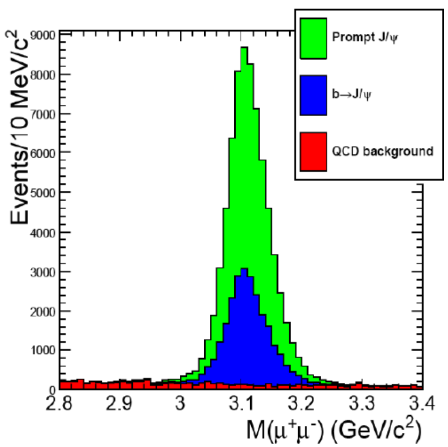
  An extra generation of ordinary fermions is excluded at the 6σ level on the basis of the S parameter alone... [Erler & Langacker]

  This result assumes that...any new families are degenerate [Erler & Langacker]

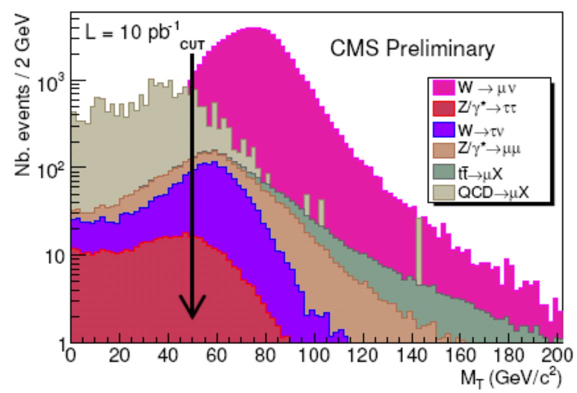
  Just as our 3rd generation??? [Holdom; Kribs, TP, Spannowsky, Tait]

Flame-bait by Tilman Plehn

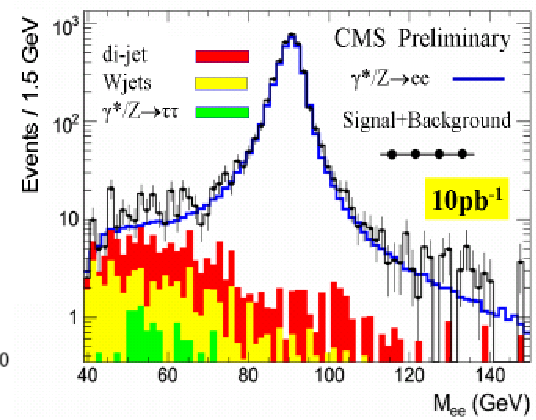
Commissioning with standard candles: “ontogeny recapitulates phylogeny”



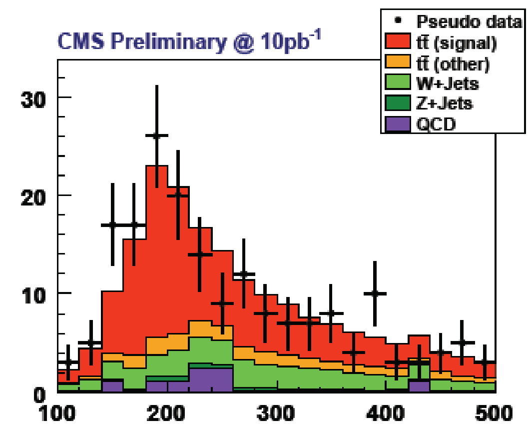
J/ψ: low p_T leptons



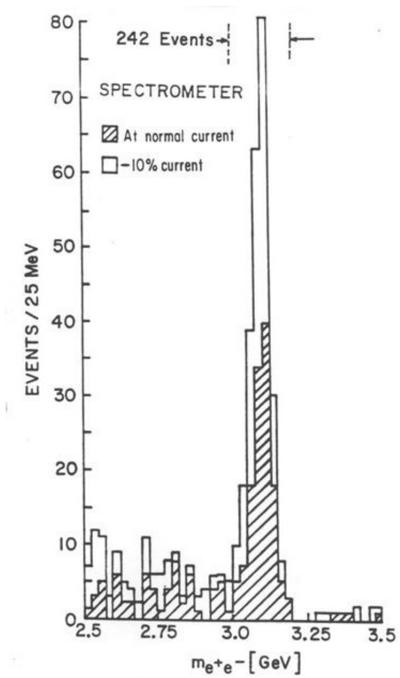
W: high p_T leptons and MET

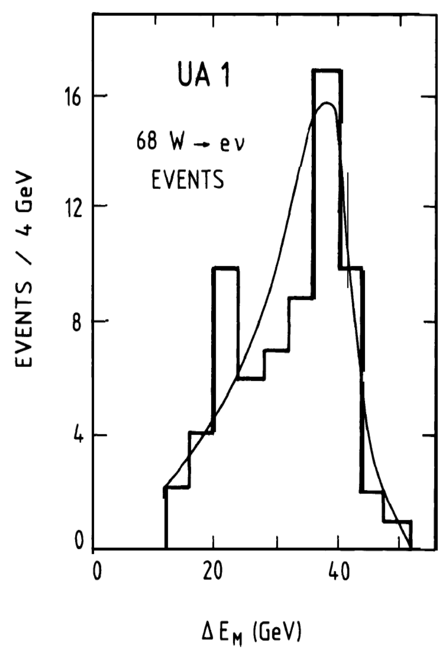


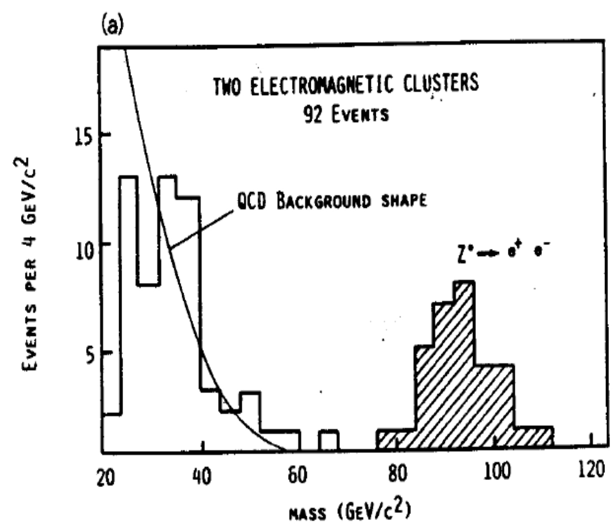
Z: high p_T leptons

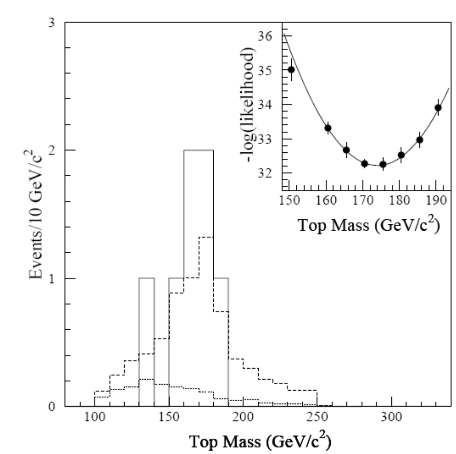


Top: high p_T leptons and jets, MET, b tagging





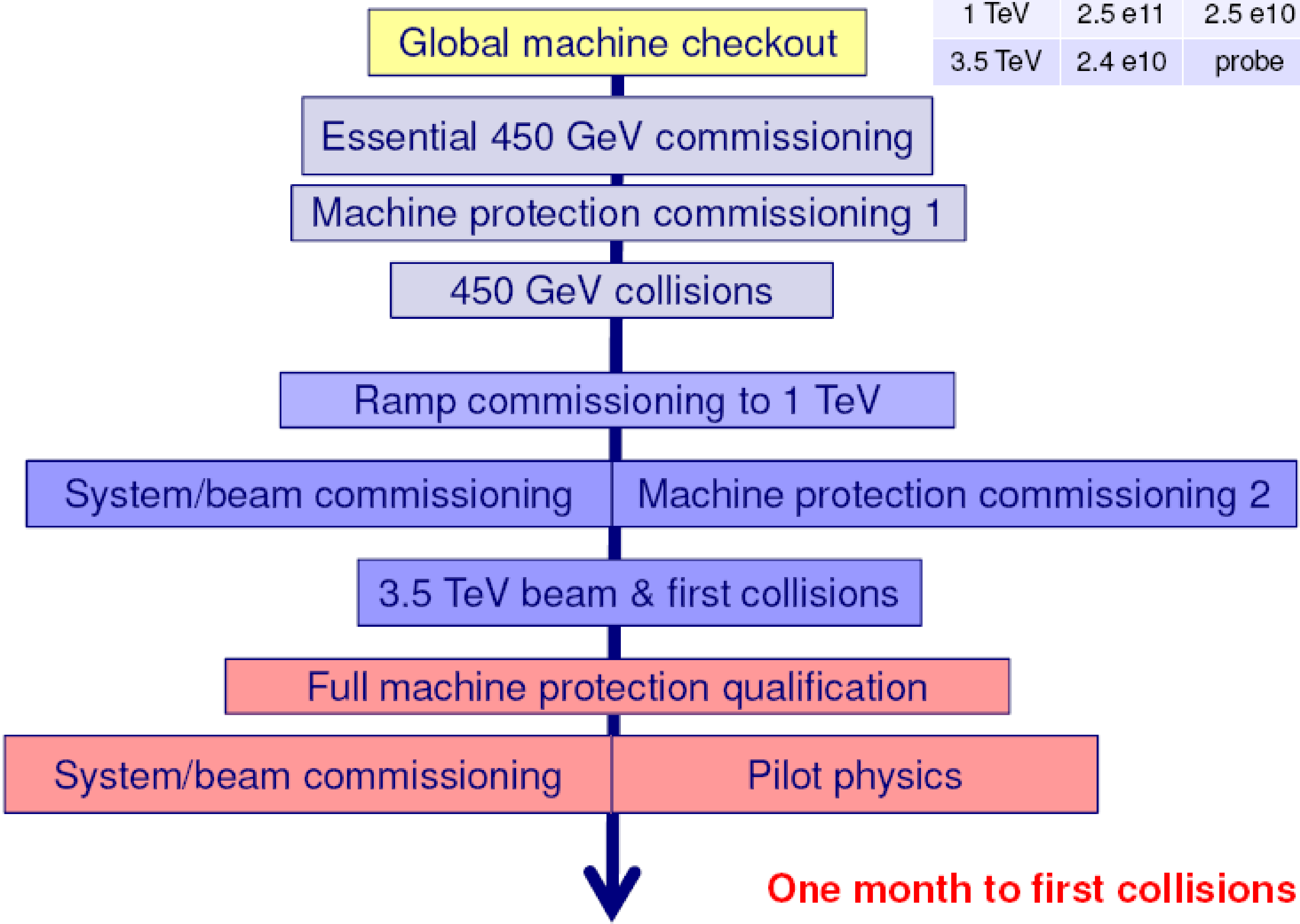




Following slides from Mike Lamont, **“LHC Status and 2009/2010 operations”**, 2009/9/7, CMS Week in Bologna

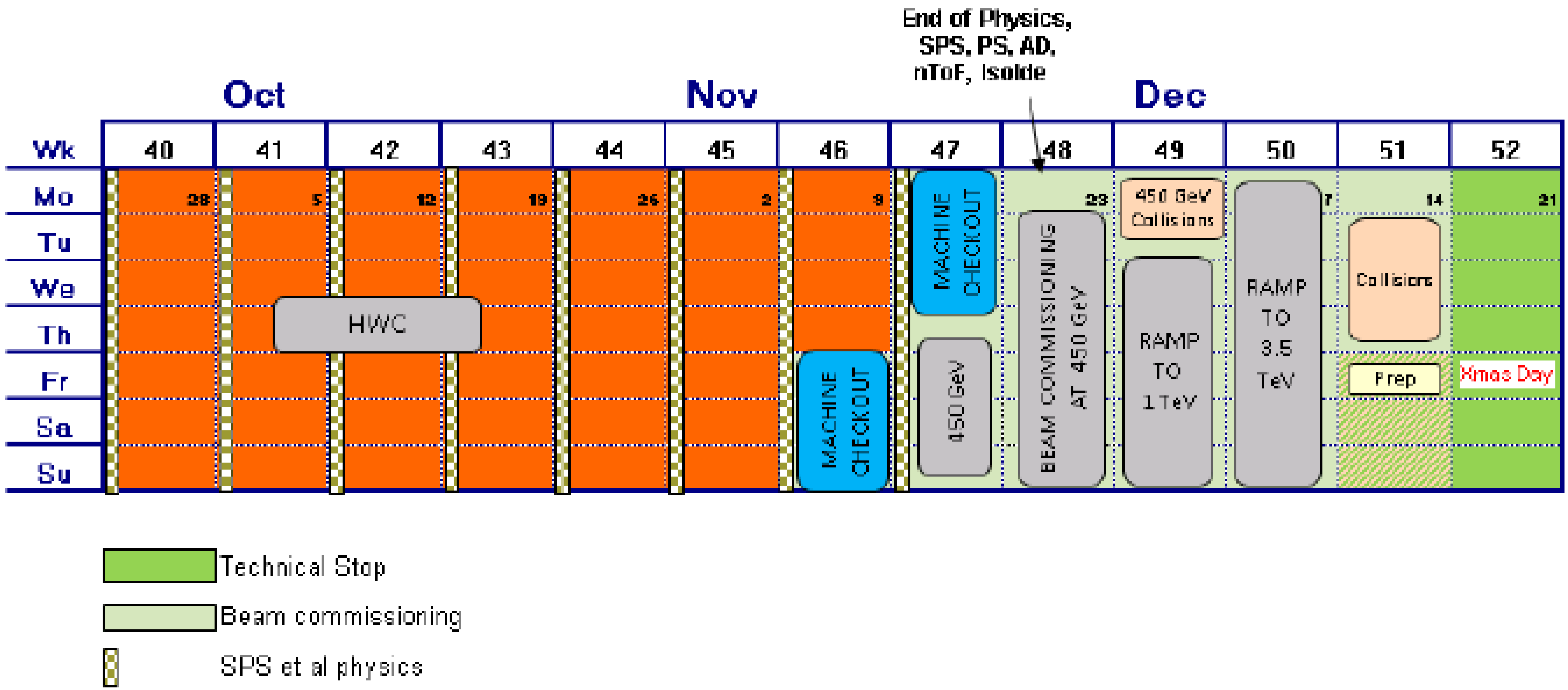
Beam commissioning

Energy	Safe	Very Safe
450	1 e12	1 e11
1 TeV	2.5 e11	2.5 e10
3.5 TeV	2.4 e10	probe



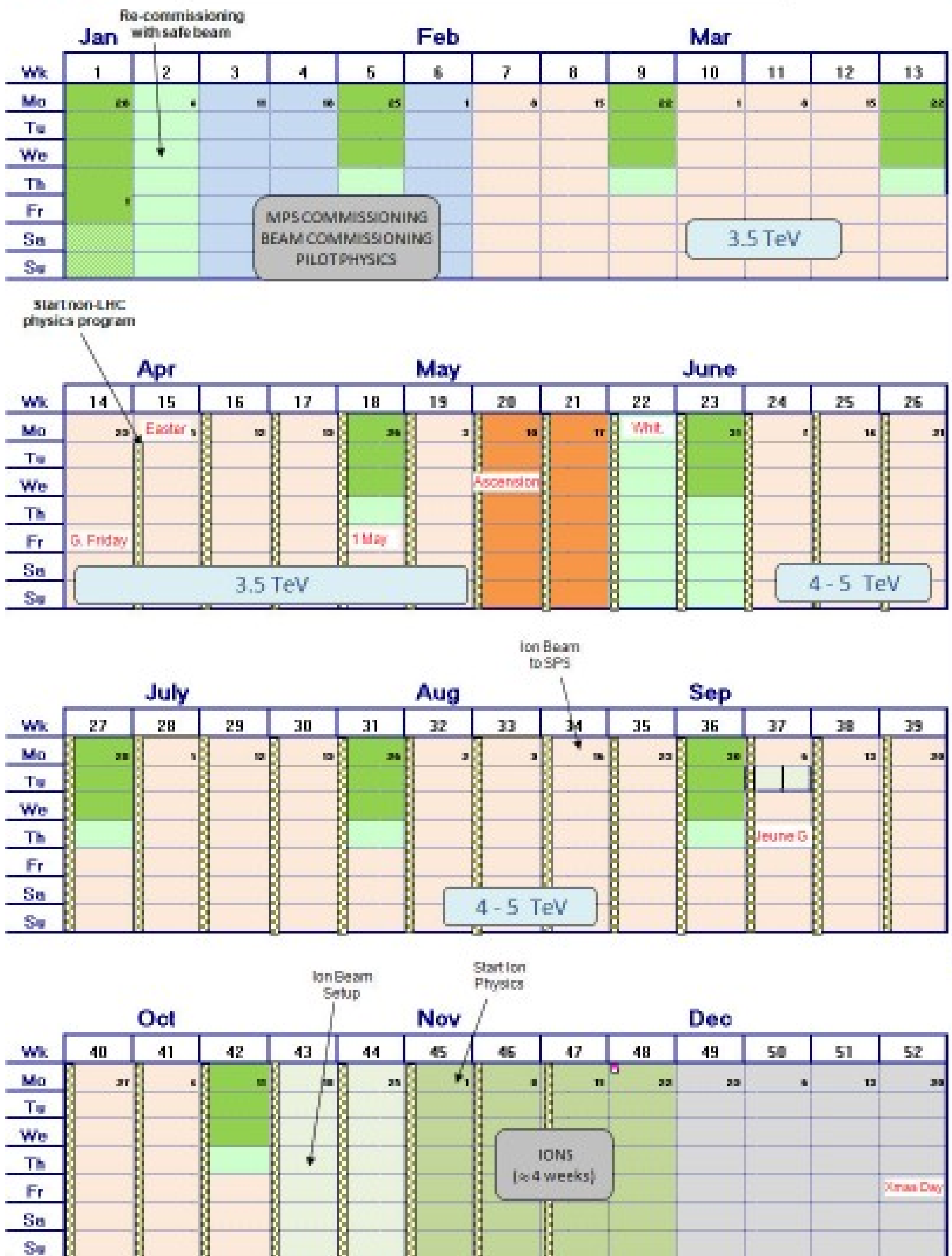
One month to first collisions

LHC 2009



- All dates approximate…
- Reasonable machine availability assumed
- Stop LHC with beam ~19th December 2009, restart ~ 4th January 2010

LHC 2010 – very draft



- **2009**:
    - 1 month commissioning
- **2010**:
    - 1 month pilot & commissioning
    - 3 month 3.5 TeV
    - 1 month step-up
    - 5 month 4 - 5 TeV
    - 1 month ions