Simulating the $H \to WW \to l\nu l\nu$ Channel

Including new theoretical developments into an experimental analysis

Anne-Sylvie Giolo-Nicollerat (CERN-CMS)

in collaboration with ETH-Zürich group

Workshop sui Monte Carlo, la Fisica e le Simulazioni a LHC

Frascati, May 22nd 2006

Outline

Goal: estimate CMS discovery potential for $H \rightarrow WW$

How to get a good simulation ? And at NLO ?

- Generation of signal and backgrounds at (N)NLO: reweighting
- Generating top background: generator comparison
- Including singly resonant top production at NLO

The H→WW→lνlν channel

First proposed Dittmar, Dreiner, Phys. Rev.D55, 1997

- Discovery channel for $2m_W<m_H<2m_Z$
- Branching ratio close to 1 in this region
- But high backgrounds and no narrow mass peak reconstruction (need good S/B and good background understanding)

Signature for H→WW→lνlν

2 leptons plus missing energy

Main backgrounds: WW, ttbar

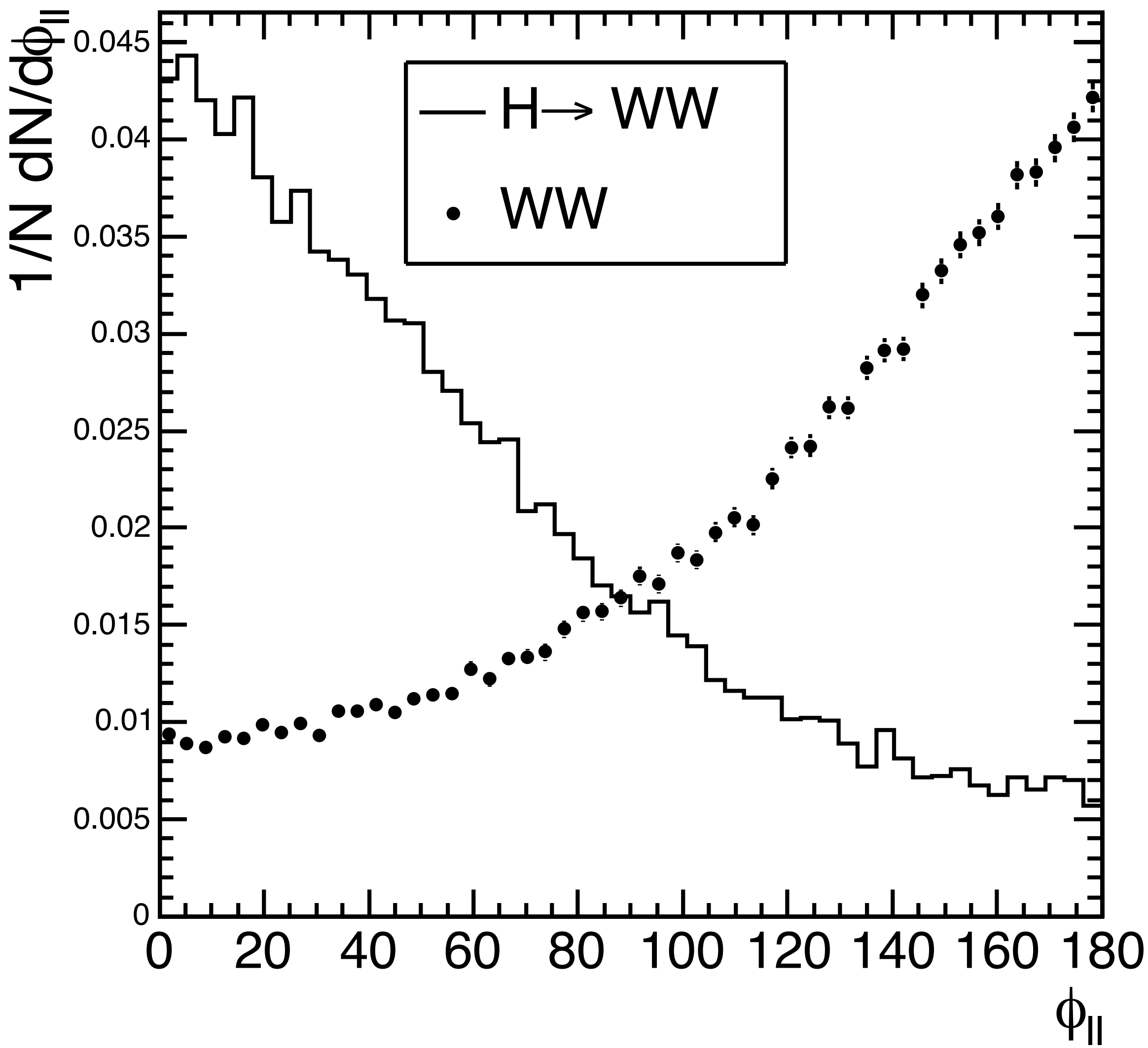
Selection based on

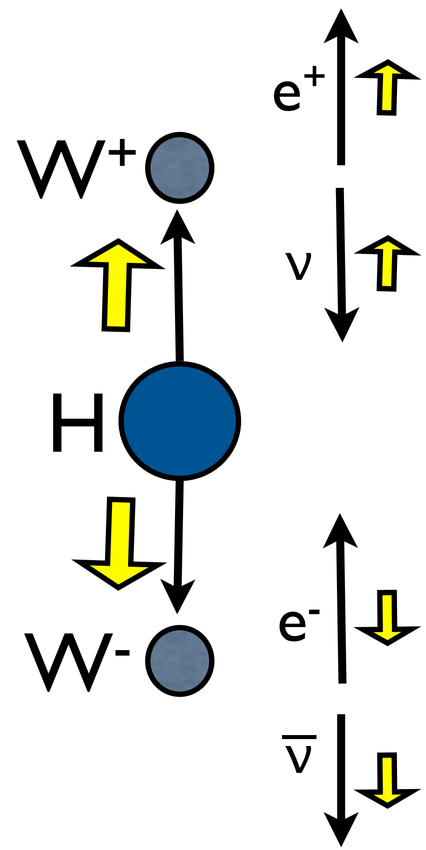
Spin correlations (reduce WW)

Jet veto (reduce ttbar)

Leptons opening angle

generated distribution





Higgs: small opening angle
WW: large opening angle

Signal and backgrounds

Cross sections times e,μ,τ branching ratio (NLO)

- Higgs 165GeV   gluon fusion: 2pb, qqH: 0.3pb

Main backgrounds

- ttbar (doubly resonant top production)   86pb
- twb (singly resonant top production)   3.4pb
- ww   12pb
- gg→ww (LO)   0.5pb

⇒ Need a good background rejection/understanding

Generation of signal and backgrounds at (N)NLO the reweighting technique

Signal and Background Generation

- need to know production cross sections and distributions of discriminating observables as precisely as possible
- Higher order QCD corrections are available
- simplest approach: inclusive K-factors

$gg \to H$	LO (PYTHIA)	NLO (MC@NLO)	NNLO (FEHIP)
σ[pb]	12.20	23.75	27.78
K_{inc}	2.28	1.18	1

- multiplication of each event with K_{inc} would give correct inclusive cross sections …
- … but not necessarily correct distributions and acceptance

A practical example:

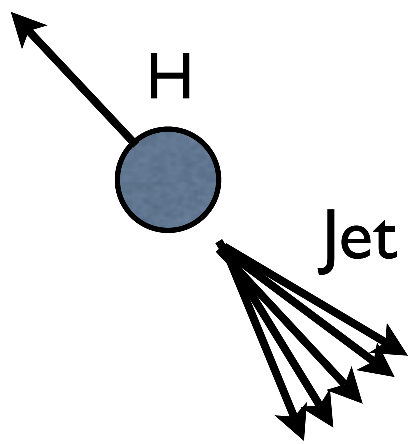
Approximating Higgs p_t spectrum using PYTHIA

$p_t(H)=0$ at LO

Need **higher order corrections** to predict **p_t spectrum**

PYTHIA LO+parton shower : mimic higher-order behavior

⇒ Higgs gets a p_t through the parton shower



Jet veto ≡ cut on the Higgs p_t spectrum

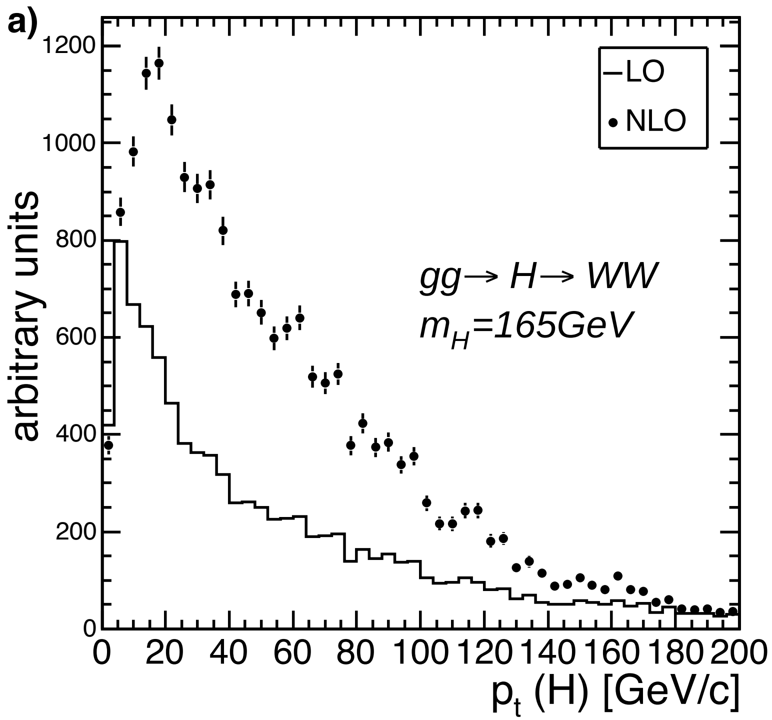
Need an accurate description of Higgs p_t spectrum

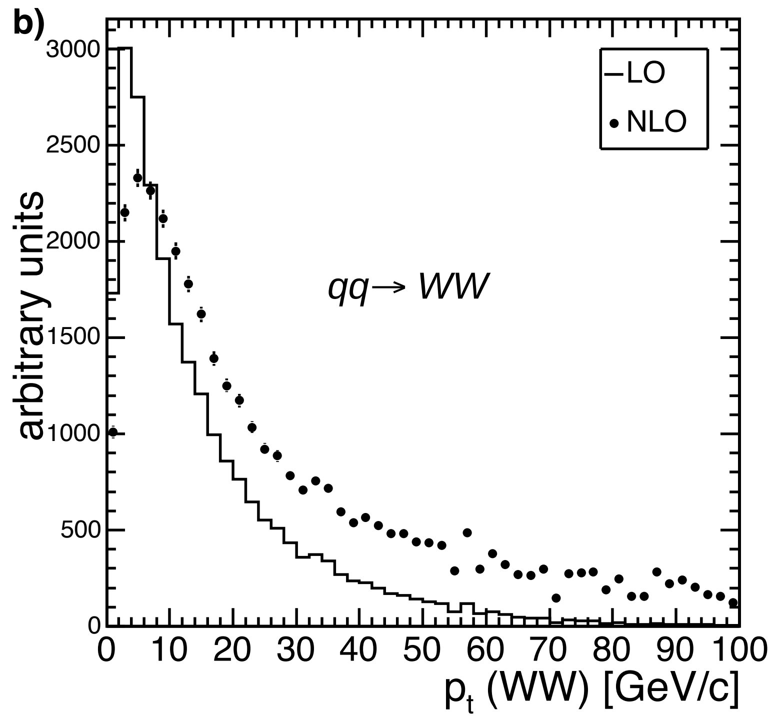
Reweighting

[Davatz et al. JHEP 0504, 2004]

p_t(WW) spectrum shape : PYTHIA ≠ NLO (MC@NLO)

reweight p_t spectrum ⇒ p_t dependent k-factors





⇒ Apply method on gg→H and WW production

Going one step further

[Davatz et al, hep-ph/0604077]

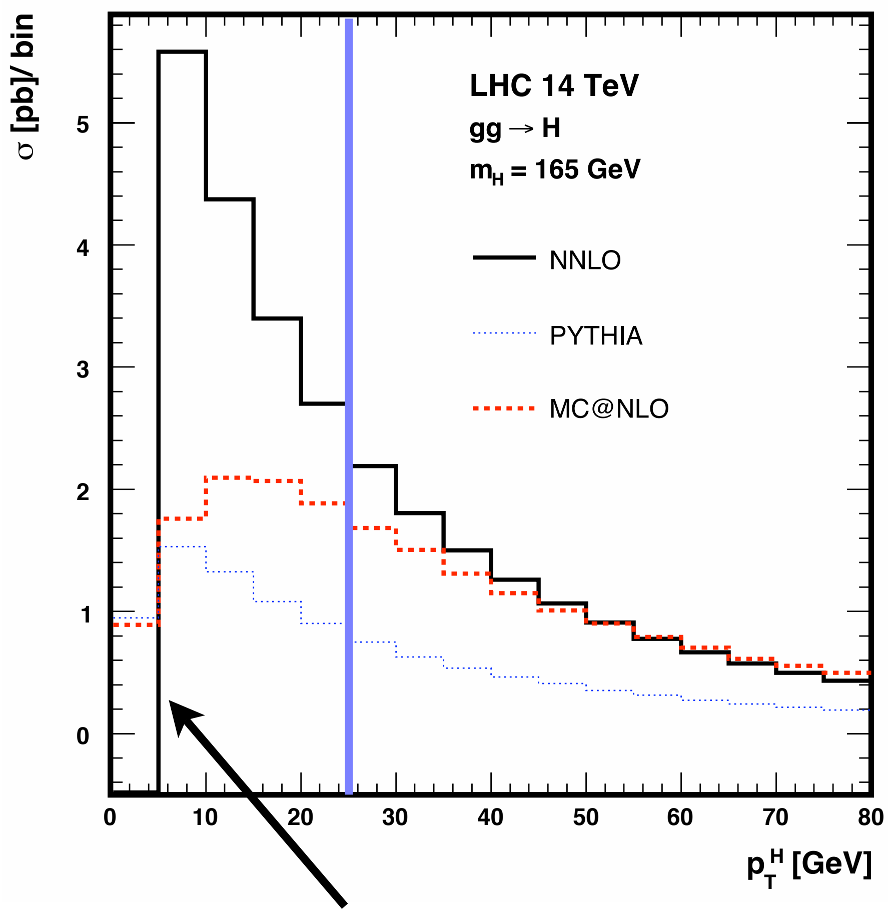
- NNLO parton generator FEHIP [Anastasiou, Melnikov, Petriello, Nucl.Phys.B724:197-246(2005)]
  $\Rightarrow$ reweight MC@NLO(3.2) and PYTHIA(6.325)
- Reweighting in p_t(H) **and** Y(H)

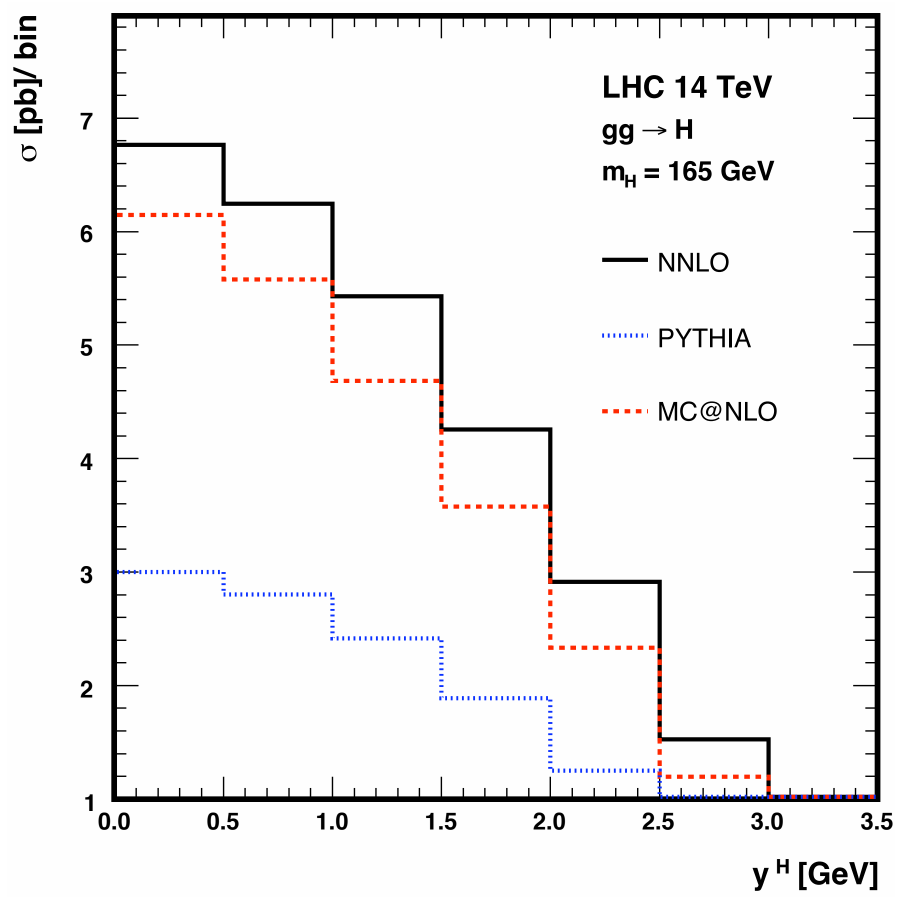
$$K_{ij}^{G} := K^{G}(\{p_f\}) = \frac{\Delta\sigma_{ij}^{\mathrm{PT}}}{\Delta\sigma_{ij}^{G}} \quad \text{if } p_\perp \in \left[p_\perp^j, p_\perp^{j+1}\right] \text{ and } Y \in \left[Y^i, Y^{i+1}\right]$$

Generators parameter set	
scales	$\mu_R = \mu_F = 1/2 \cdot m_H$
PDF	MRST2001 LO/NLO/NNLO
Higgs mass	165 GeV

☛ Full event kinematic determined through p_t, Y and jet multiplicity

Higgs transverse mass and rapidity

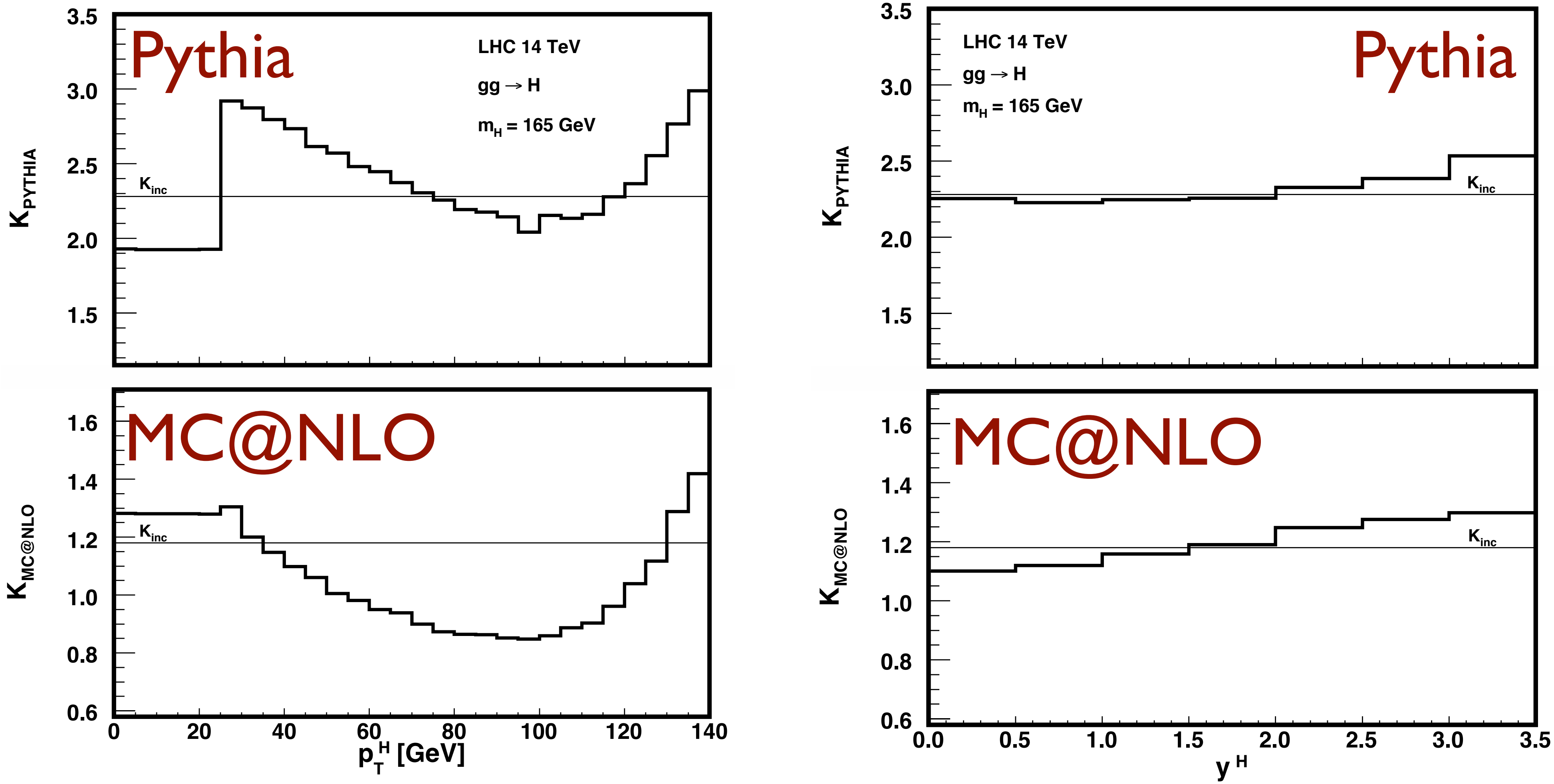




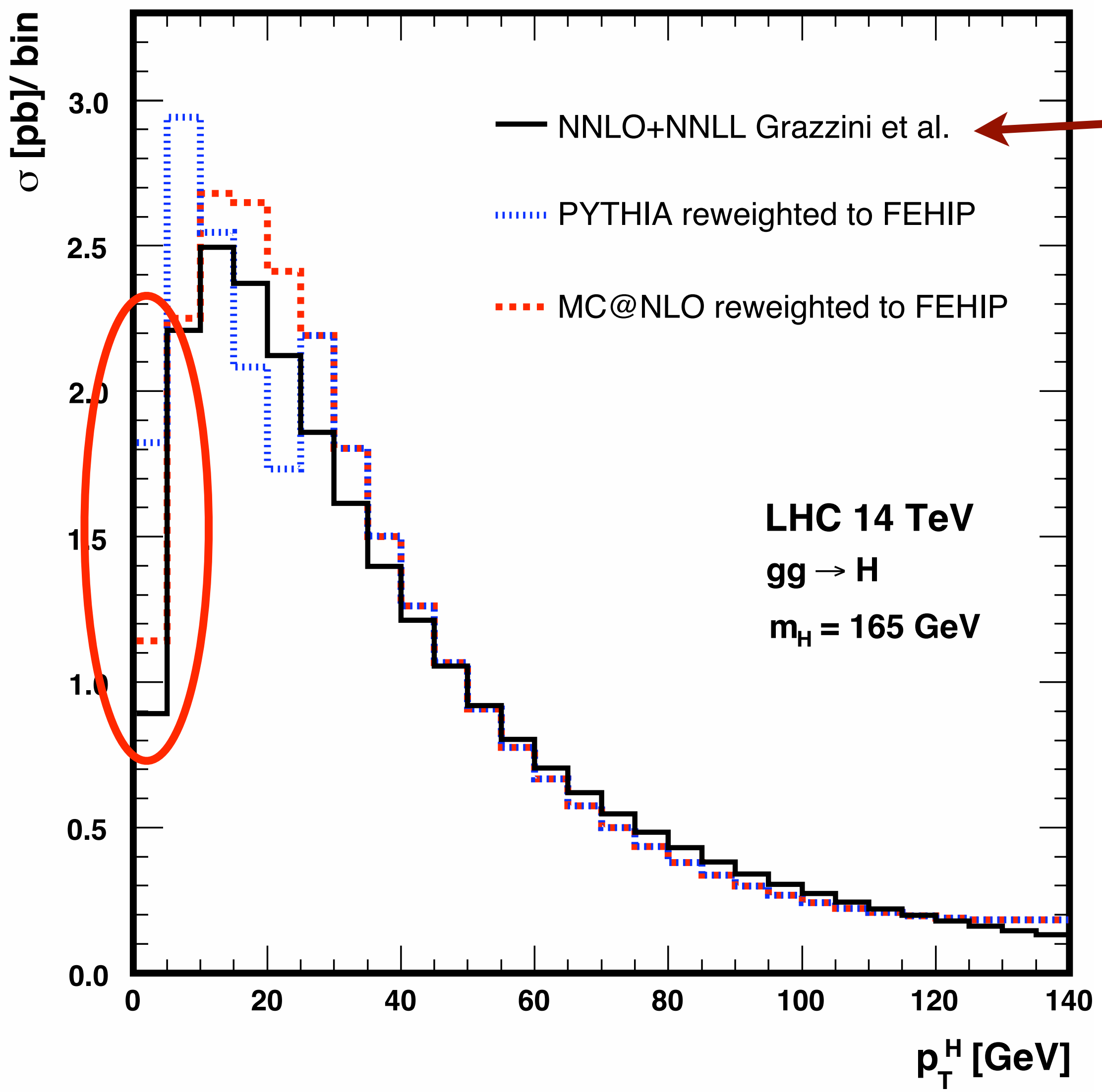
In Fehip, no log resummation
Reweighting: constant K-factor from 0 to 25GeV
rely on parton shower prediction for this region

The K-factors

Reweight Pythia and MC@NLO to FEHIP



Low p_t(H) description



Compare with calculation including **logarithm resummation** (NNLO+NNLL) Grazzini et al. [hep-ph/0302104] [hep-ph/0508068]

Low p_t region: Still some uncertainty due to different models used

Jet veto efficiency

- Veto jets with E_T>30GeV (H→WW selection cuts)
- Jet veto efficiency (NNLO)

  0.47 Pythia rwth. to $FEHIP^{NNLO}$
  0.54 MC@NLO rwth. to $FEHIP^{NNLO}$
  0.54 FEHIP NNLO

  6% spread

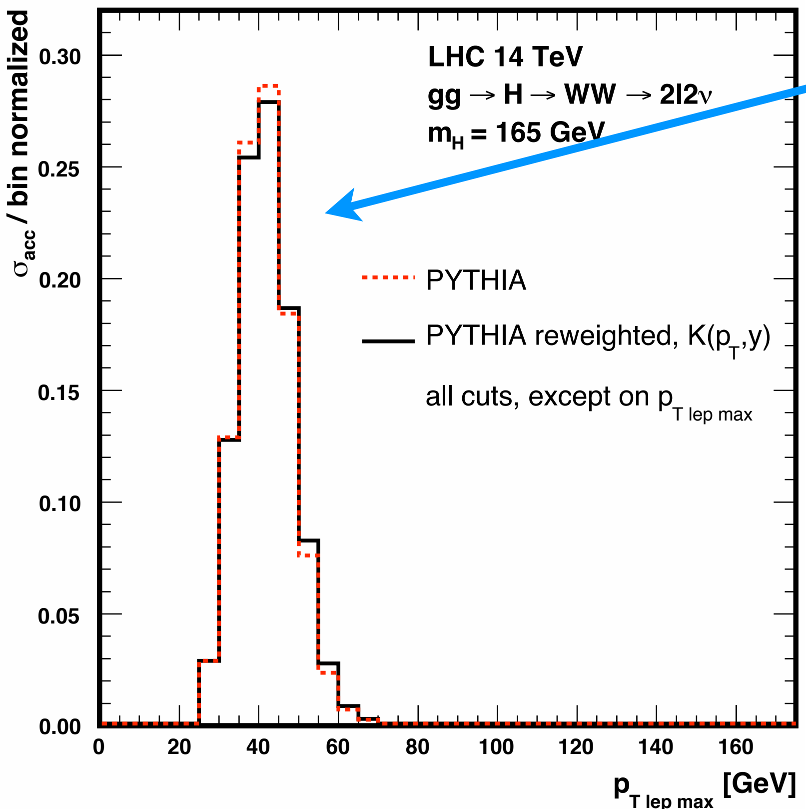
- Comparison: jet veto efficiency (NLO)

  0.50 Pythia
  0.51 MC@NLO
  0.60 Pythia rwth. to $FEHIP^{NLO}$
  0.68 FEHIP NLO

  30% spread

At NNLO, jet multiplicity is better described
⇒ leads to smaller discrepancy between predictions

Kinematic variables



p_t(lep): very similar

Apply selection cuts and (pythia) reweighting on:

p_t → global K factor of 2.0

p_t and Y → global K factor of 2.1

⇒ Reweighting to Y <5% effect

Next step: add in FEHIP W decays to study lepton angles

Reweighting is a good phenomenological approach to include higher order corrections in event generators

Simulation of top background: generator comparison

Simulating ttbar background

[Davatz, Zanetti, A.G.]

High cross section, need reduction of $O(10^5)$...

What is the accuracy of Monte Carlo predictions ?

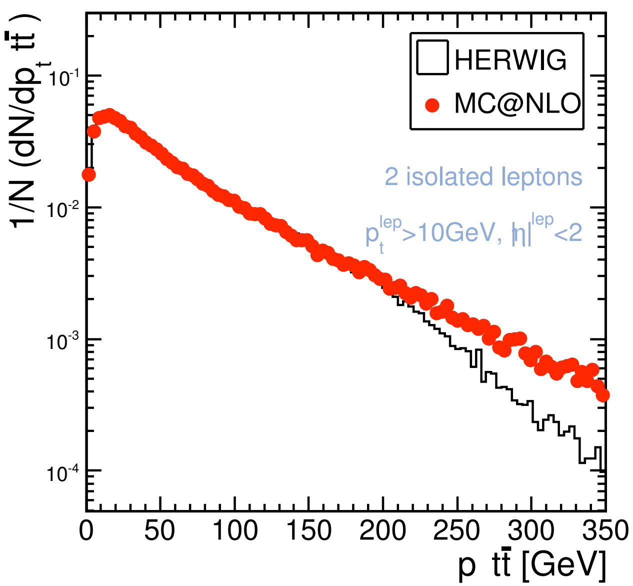
Compare the predictions of
PYTHIA, TOPREX, HERWIG and MC@NLO

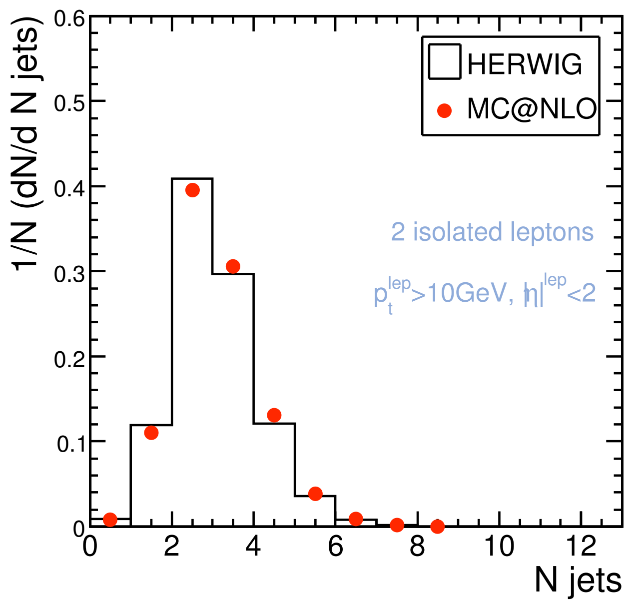
	PYTHIA 6.227	TopREX 4.11	HERWIG 6.508	MC@NLO 2.31
Matrix Elements	LO	LO	LO	NLO
hadronization model	LUND	LUND	Cluster	Cluster
shower model	Q^2 ordered	Q^2 ordered	angular ordered	angular ordered
spin correlations between t and $\bar{t}$	no	yes	yes	no

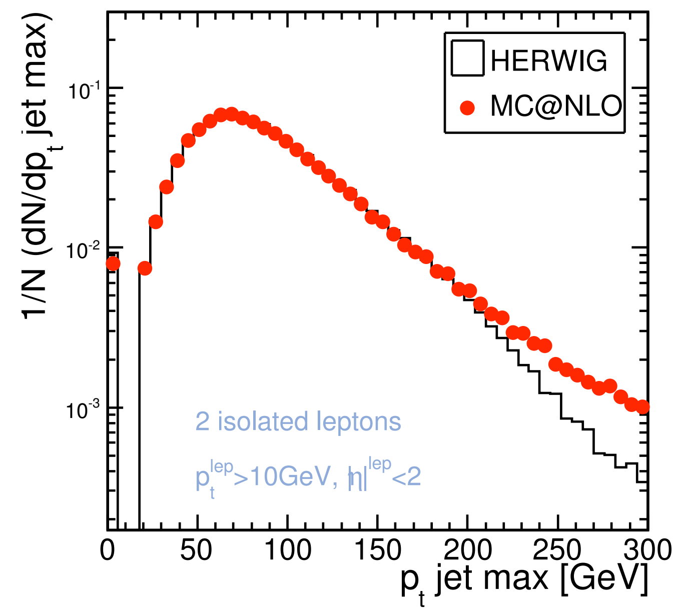
study on generator tree (add cuts to include CMS acceptance)

ttbar: do we need reweighting?

Compare HERWIG and MC@NLO







Small (<10%) difference in jet veto efficiency

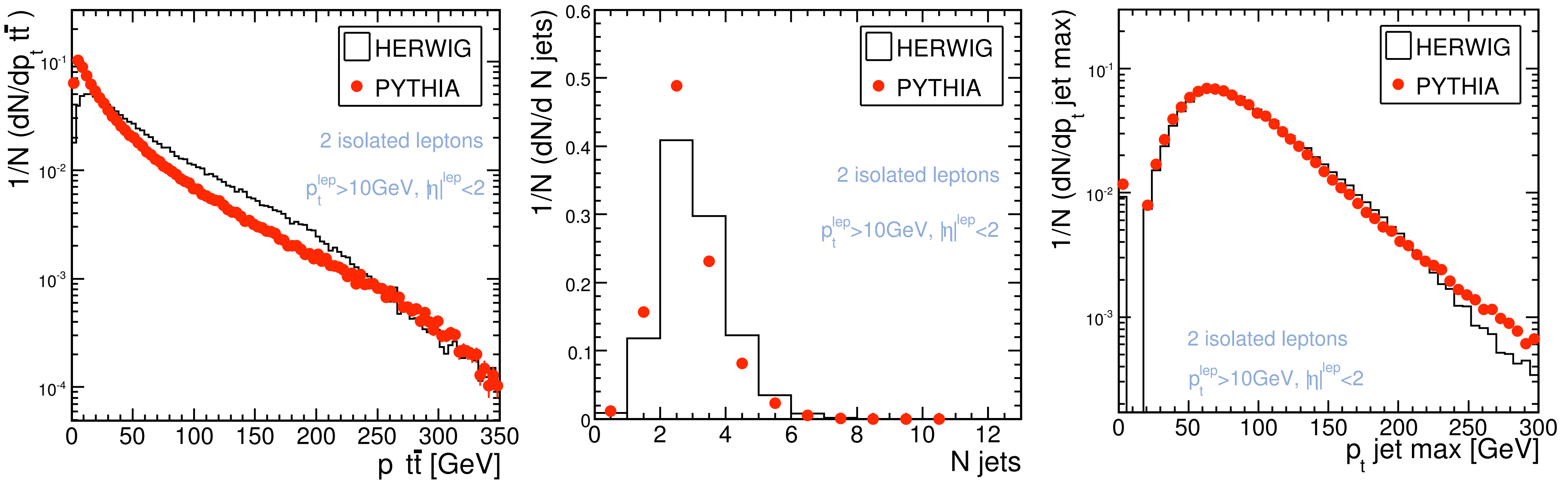
(already two hard jets in LO)

Similar lepton kinematics

⇒constant k-factors can be used

ttbar: effect from the different shower models

Compare PYTHIA and HERWIG



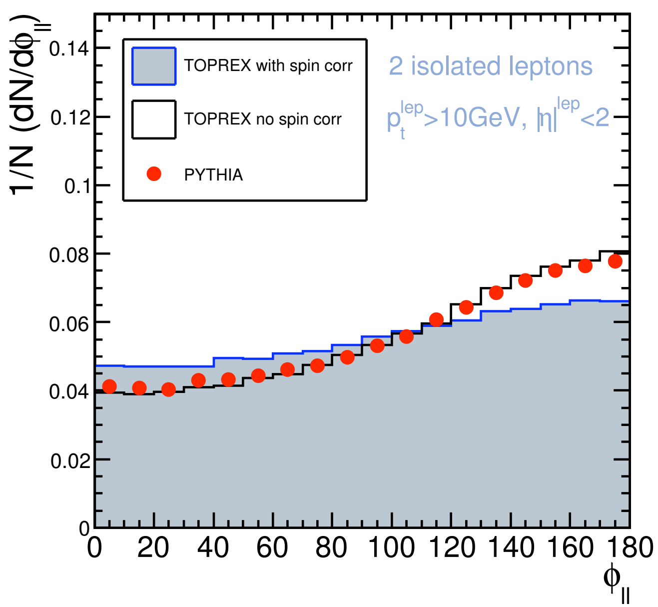
Pythia produces less and softer jets

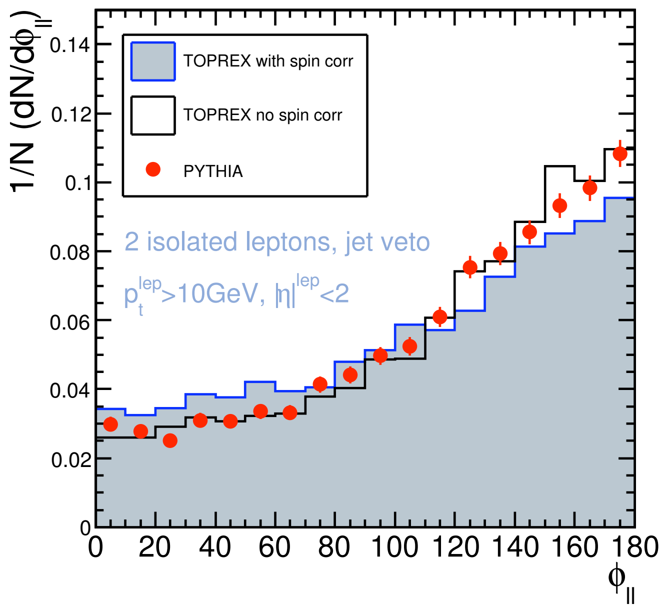
Sizable difference in jet veto efficiency (~15%)

Other cut efficiencies similar

ttbar: spin correlations

Compare PYTHIA and TOPREX





Including spin correlation flattens Φ_{ll} distribution
(but Φ_{ll} more similar after jet veto)

Difference of ~10% in selection efficiency

ttbar generator study - conclusion

Cannot rely only on ttbar Monte Carlo predictions

(up to 30% spread in the predictions...)

Differences can be constrained/understood from data

Use a normalization region to estimate ttbar in the signal region

$$N_{\text{bkg}}^{\text{signal region}} = \frac{\varepsilon_{\text{bkg}}^{\text{signal region}}}{\varepsilon_{\text{bkg}}^{\text{normalization region}}} N_{\text{bkg}}^{\text{normalization region}}$$

see M. Zanetti's talk tomorrow

Theoretical errors are also reduced in the efficiency ratio

Vary renormalization and factorization scale :

40-50% variation in total cross section at LO **5%** variation in the ratio !

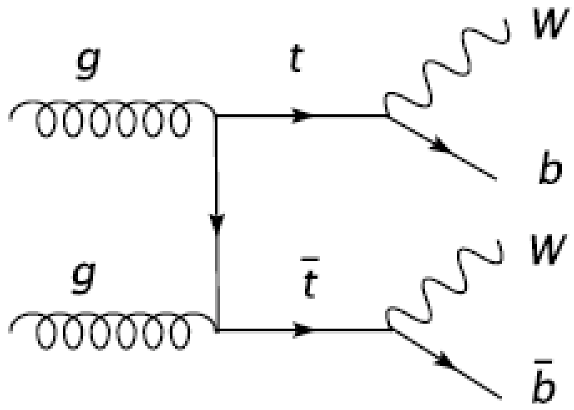
[N. Kauer, Phys.Rev.D70, 2004]

Simulating singly resonant top production

Simulating singly resonant top

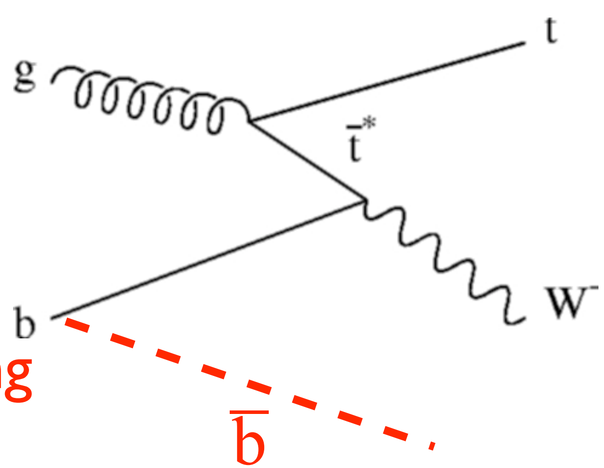
[solution proposed in les Houches '05]

~ 90% of total top cross section is doubly resonant ttbar



~10% one off-shell top (twb)

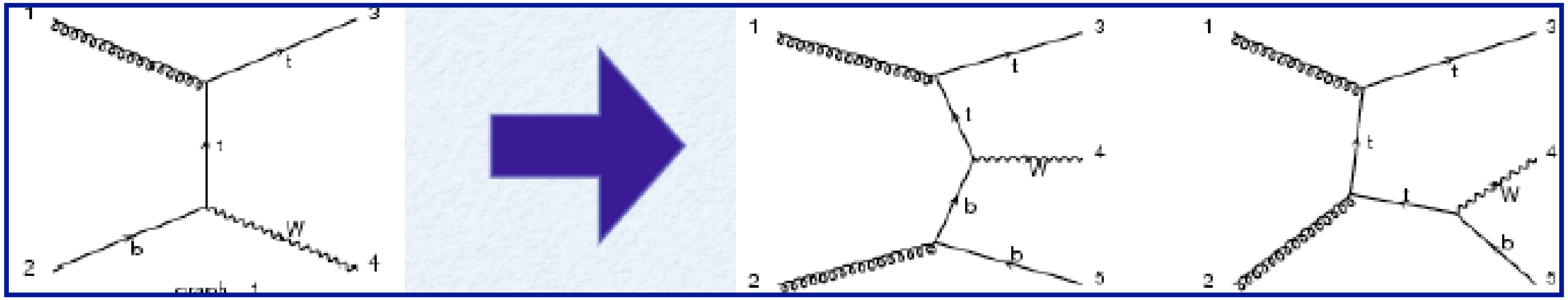
This fraction increases after jet veto



Need to consider also singly resonant top production

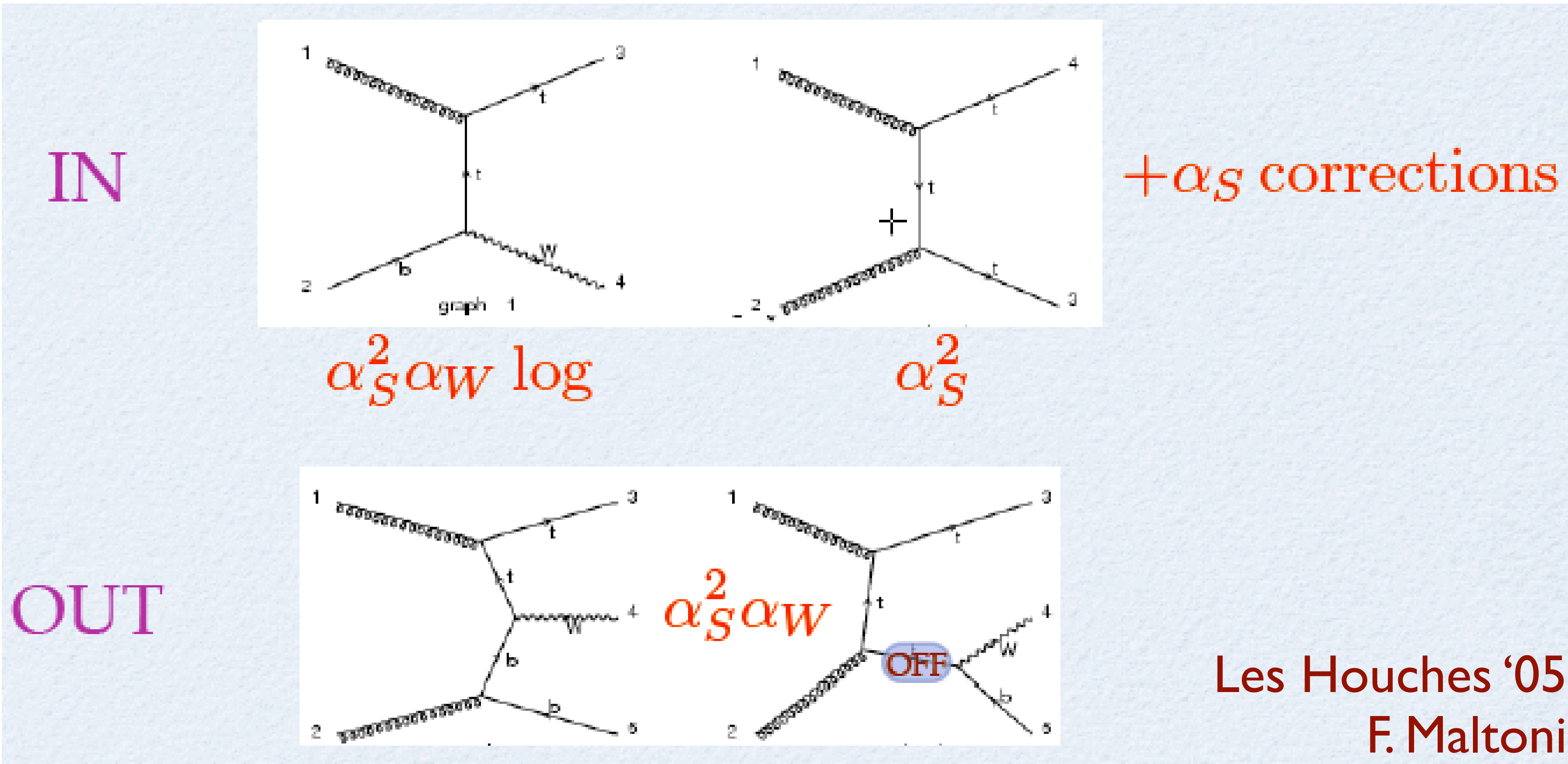
How to get a NLO estimate of twb ?

Problem: NLO twb contains LO ttbar !!



ttbar and twb are well separated when a jet veto is applied
[Campbell, Tramontano, hep-ph/0506289]

Leave out diagrams proportional to $\alpha^2_s\alpha_w$



- ✦ Interference diagrams are neglected
- ✦ Gauge invariant solution
- ✦ Physically equivalent to a jet veto !
- ✦ But k-factor for twb is phase space dependent

Is twb kinematic different at NLO ?

can we use a constant k-factor ?

Compare MCFM NLO and LO (parton level)
with TOPREX (parton shower from PYTHIA)

	MCFM				TopREX
	LO		NLO		LO
Selection cuts	$\sigma \times BR$ (fb)	rel. eff	$\sigma \times BR$ (fb)	rel. eff	rel. eff
No cuts	271		377		
2 lep, $\lvert\eta\rvert < 2$, $p_t > 20$GeV	204	0.75	277	0.73	
$E_t^{miss} > 40$	148	0.73	209	0.75	0.75
$\phi_{\ell\ell} < 45$	20.8	0.14	34.4	0.16	0.17
5 GeV $< m_{\ell\ell} <$ 40 GeV	10.6	0.51	15.6	0.45	0.50
Partonic jet veto, 40 GeV	1.55	0.15	1.12	0.07	0.16
30 GeV$< p_t^{\ell\, max} <$55 GeV	1.08	0.70	0.73	0.65	0.63
$p_t^{\ell\, min} >$25 GeV	0.73	0.68	0.49	0.67	0.67

At NLO only jet veto has a different efficiency (as expected !)

In short: use parton shower MC to simulate twb and add a constant k-factor
(calculated for the region defined by signal cuts)

CMS discovery potential in the H→WW→lνlν channel

Generators used

- Signal PYTHIA
- ww PYTHIA

(Signal and ww → Reweighted with pt-dependent k-factors)

- wz, zz PYTHIA constant k-factors
- gg→ww private implementation from N. Kauer et al.
- ttbar, twb TOPREX k-factors of resp. 1.6/0.7

Plug them in a full CMS detector simulation

Selecting Higgs events

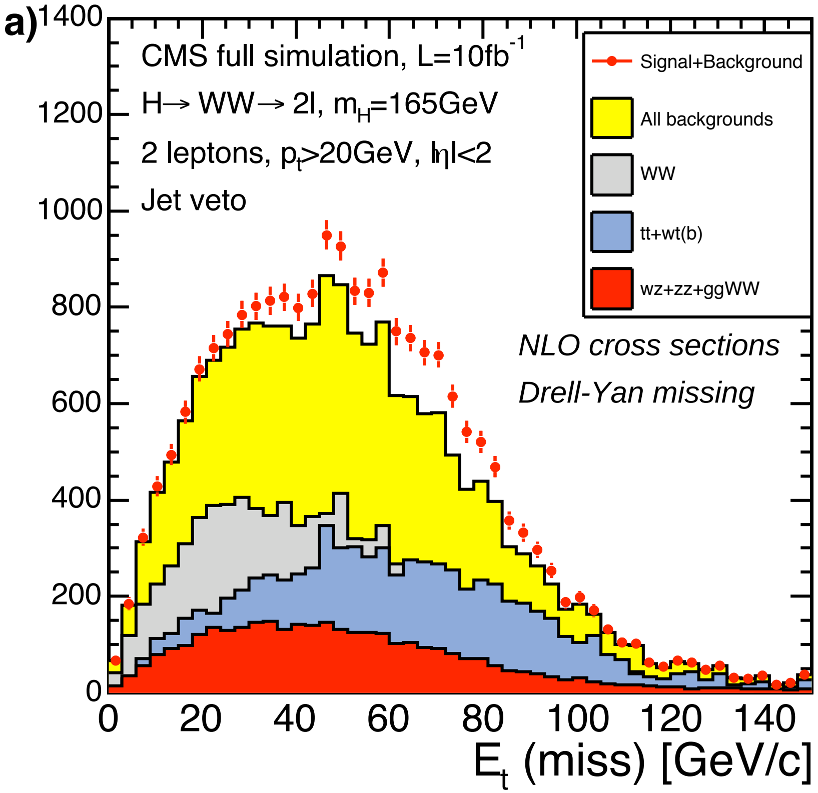
Start with 2 leptons, opposite charge:

- E_t > 20 GeV, $|\eta|<2$, $\sigma_{IP}<3$
- Missing energy > 50 GeV
- No jet with E_t^{raw} > 15 GeV and $|\eta|<2.5$
- $\Phi_{ll} < 45°$
- 12 GeV< m_{ll} < 40 GeV
- 30 GeV < $p_t(lep)^{max}$ < 55 GeV
- $p_t(lep)^{min}$ > 25 GeV

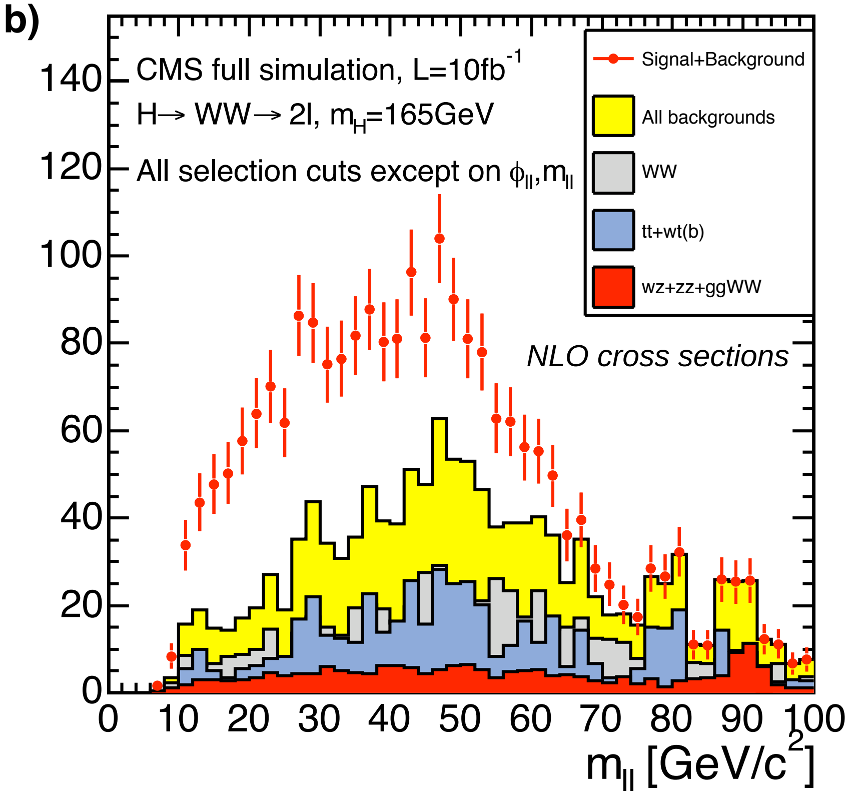
Experimental details and issues: see M. Zanetti's talk tomorrow

Some distributions
(full CMS detector simulation)

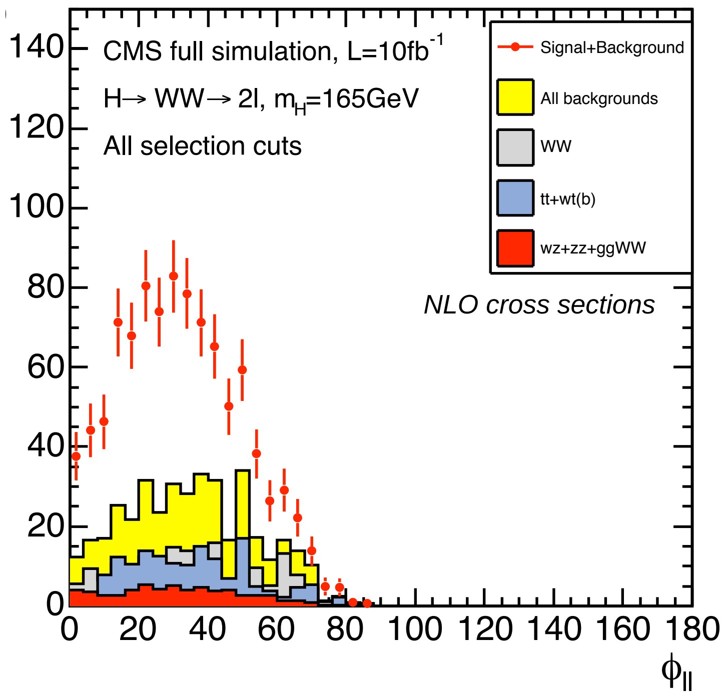
Missing energy



Leptons invariant mass



Expected signal and background



Leptons angular distribution after all cuts

Reaction $pp \to X$	$\sigma_{NLO} \times BR$	L1+HLT	2 leptons	All cuts
$\ell = e, \mu, \tau$	pb	Number of events per fb^{-1}		
$H \to WW \to \ell\ell$, $m_H = 160$ GeV	2.34	1353	359	42
$H \to WW \to \ell\ell$, $m_H = 165$ GeV	2.36	1390	393	46
$H \to WW \to \ell\ell$, $m_H = 170$ GeV	2.26	1350	376	33
$qq \to WW \to \ell\ell$	11.7	6040	1400	12
$gg \to WW \to \ell\ell$	0.48	286	73	3.7
$tt \to WWbb \to \ell\ell$	86.2	57400	15700	9.8
$tWb \to WWb(b) \to \ell\ell$	3.4	2320	676	1.4
$ZW \to \ell\ell\ell$	1.6	1062	247	0.50
$ZZ \to \ell\ell, \nu\nu$	1.5	485	163	0.35
Sum backgrounds	105	67600	18300	28

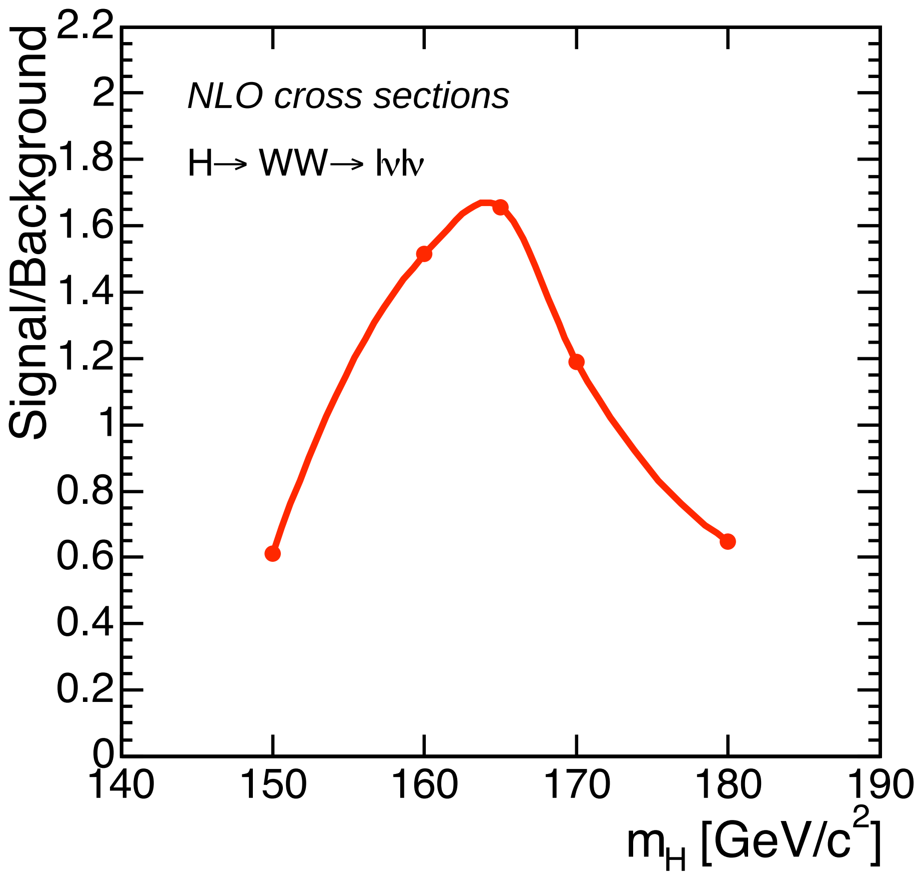
good S/B !

Assume systematics of 13% for 5fb^{-1}

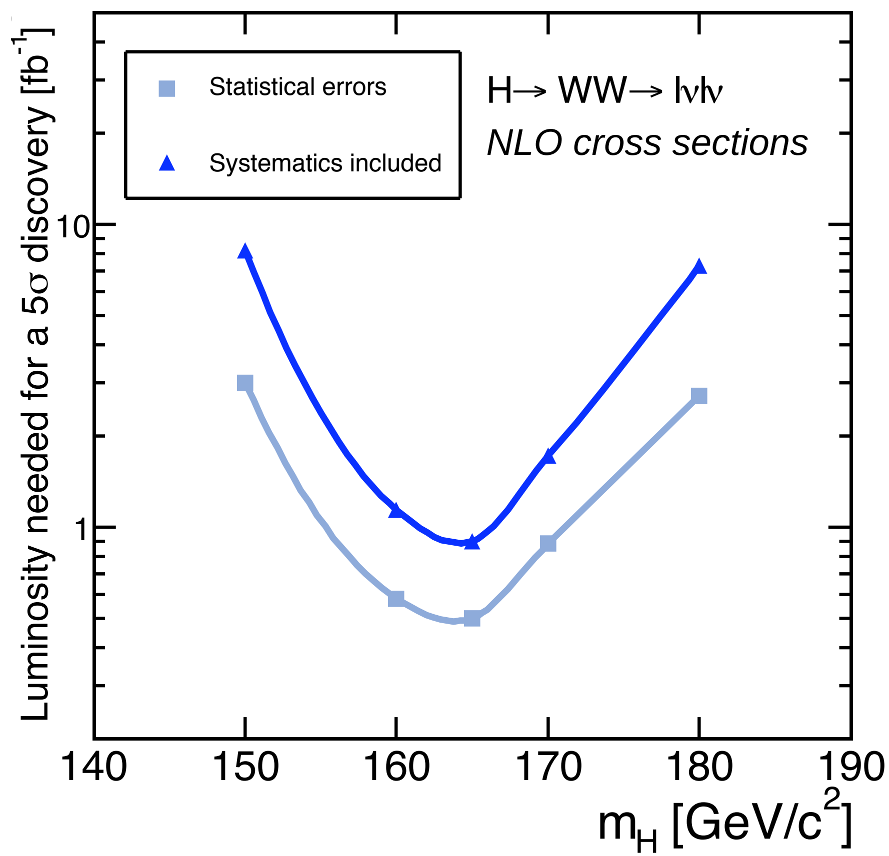
[All NLO cross sections from M. Spira and J. Campbell]

CMS Higgs discovery potential in $H \to WW \to l\nu l\nu$

Signal over background ratio



Luminosity needed for 5σ significance



Summary

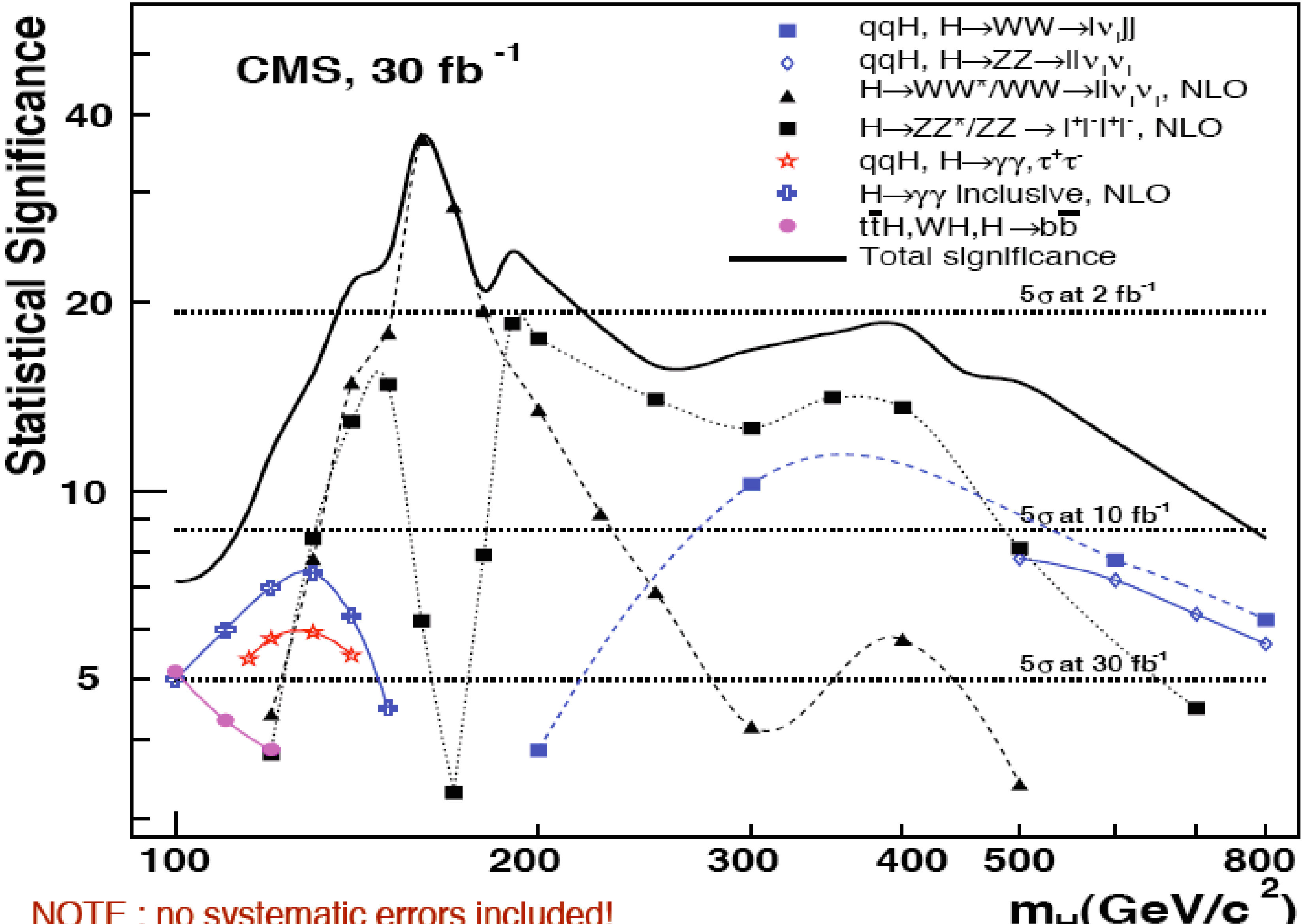
- A Higgs could be **discovered** at CMS in the WW channel with $10fb^{-1}$ if it has a mass between 150 and 180GeV
- The Higgs p_t spectrum was **reweighted** to get a more accurate estimate of signal and WW background
  - A further step is to introduce an additional reweighting in Y, leading to a <5% correction
- **Top background** prediction vary a lot from generator to another. Need background estimation from data.
- **Singly resonant top production** was simulated and represent 13% of total top background

Special thanks to G. Davatz, G. Dissertori and F. Stöckli

References

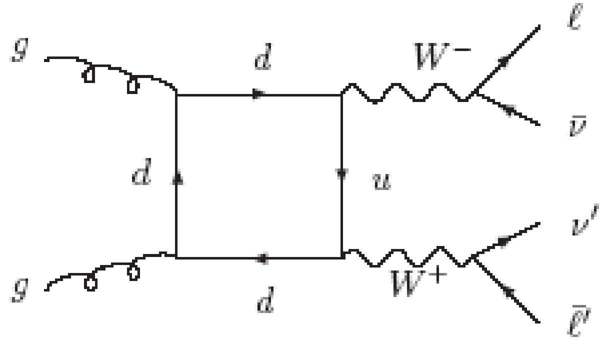
- Combining Monte Carlo generators with next-to-leading order calculations: event reweighting for Higgs boson production at LHC, Davatz et al., hep-ph/0604077
- Standard Model Higgs Discovery Potential of CMS in the $H \rightarrow WW \rightarrow l\nu l\nu$ Channel, Davatz, Dittmar, Giolo-Nicollerat, CERN-CMS-NOTE 2006/047
- Systematic uncertainties of the top background in the $H \rightarrow WW$ channel, Davatz, Giolo-Nicollerat, Zanetti CERN-CMS-NOTE 2006/048

BACKUP slides



NOTE : no systematic errors included!

Adding gluon fusion to WW production

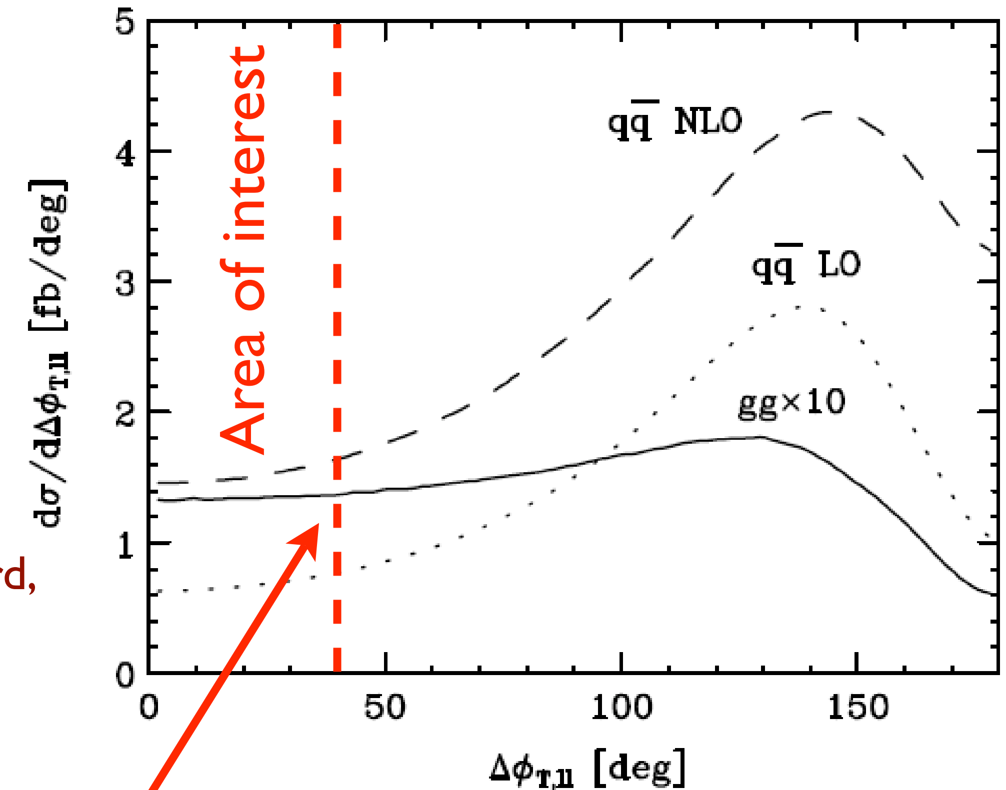


Calculated last year !

[T. Binoth, M. Ciccolini, N. Kauer, M. Kramer, JHEP 0503,2005]

[M. Duhrssen, K. Jakobs, J.J. van der Bij, P. Marquard, JHEP 0505,2005]

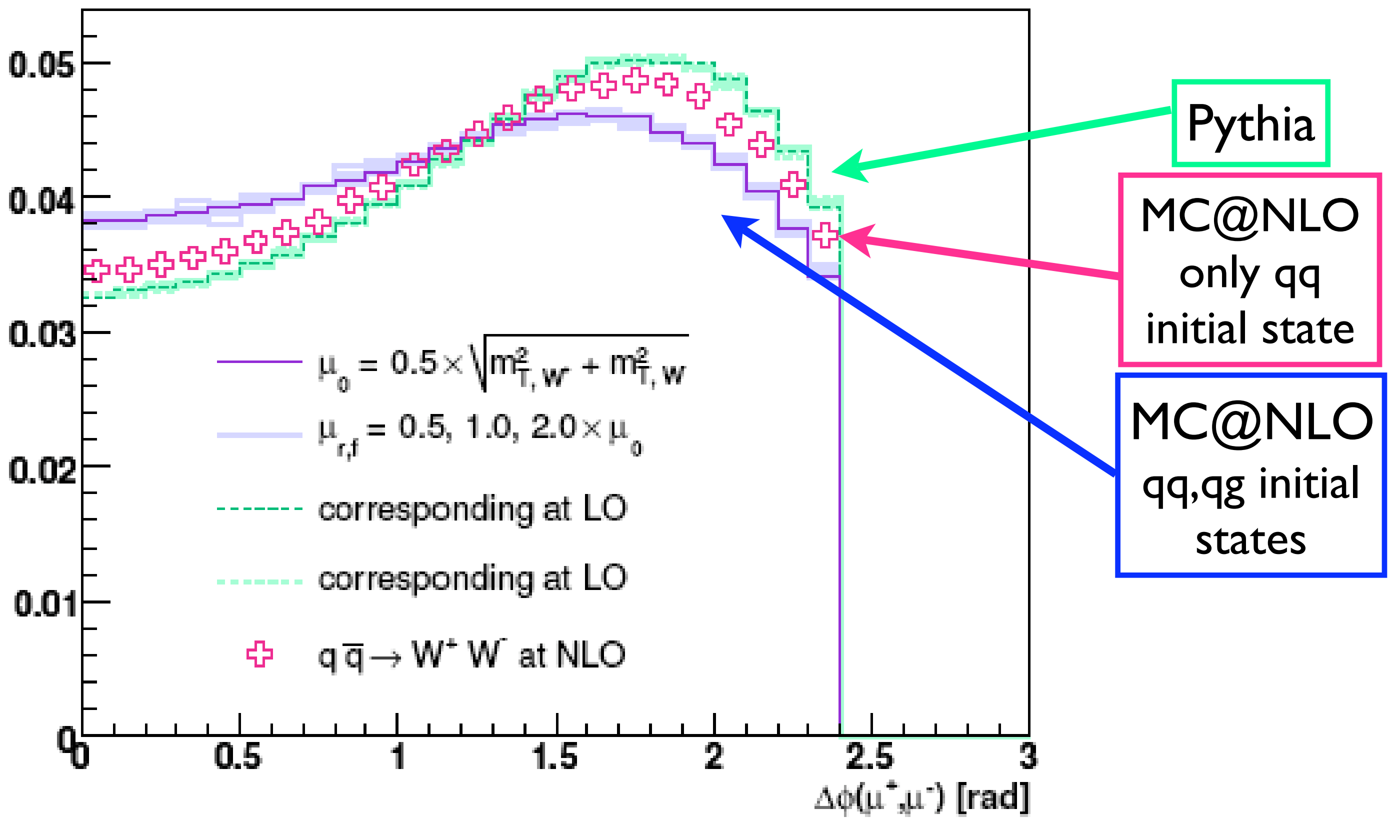
Inclusive: $$\frac{\sigma_{gg\rightarrow WW}}{\sigma_{qq\rightarrow WW}} = 4\%$$



Fraction increased by Higgs selection cuts !

A limitation:

qq→WW Φ_{ll} distribution
for MC@NLO and Pythia (reweighted on MC@NLO)



[Drollinger et al. CERN-CMS-NOTE-2005-024]

Reconstructing jets

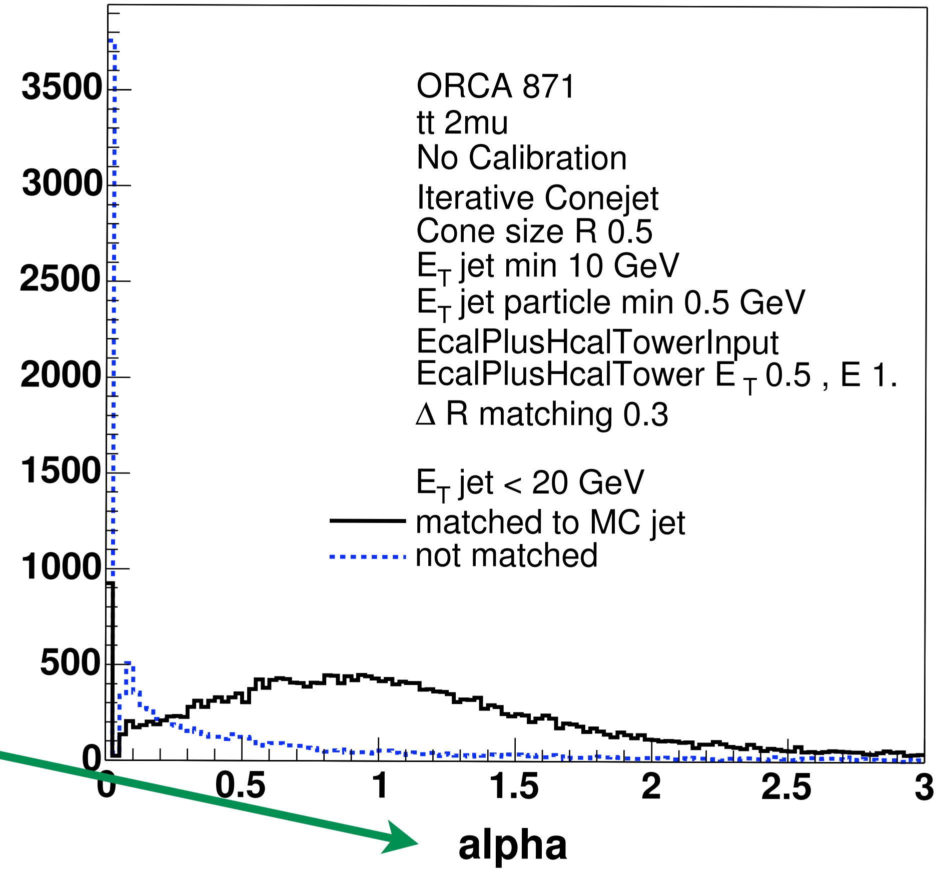
Efficiency of jet veto depends on jet reconstruction !

Cone algorithm of size 0.5

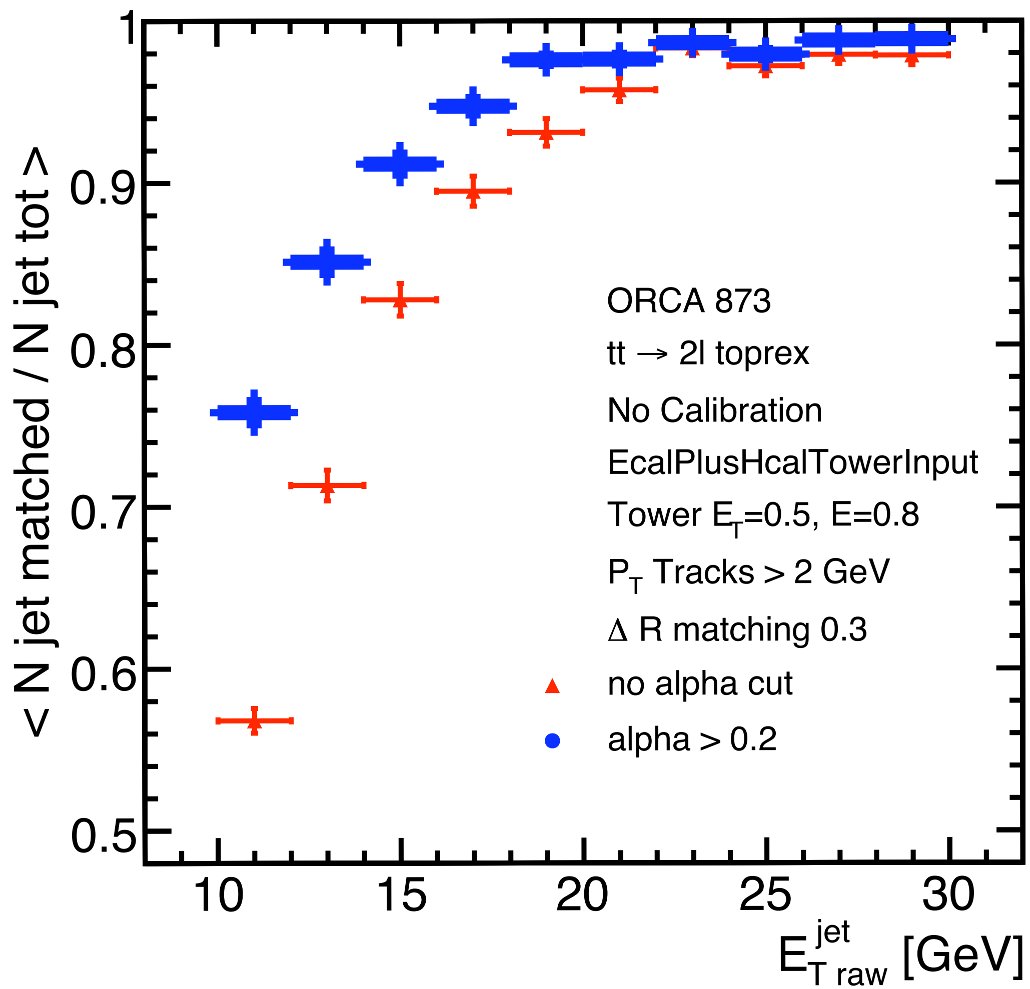
Use jet track content to reduce fake jets from noise and underlying event:

Require

$\Sigma p_t(\text{tracks inside jet})/E_{jet}^{raw} > 0.2$

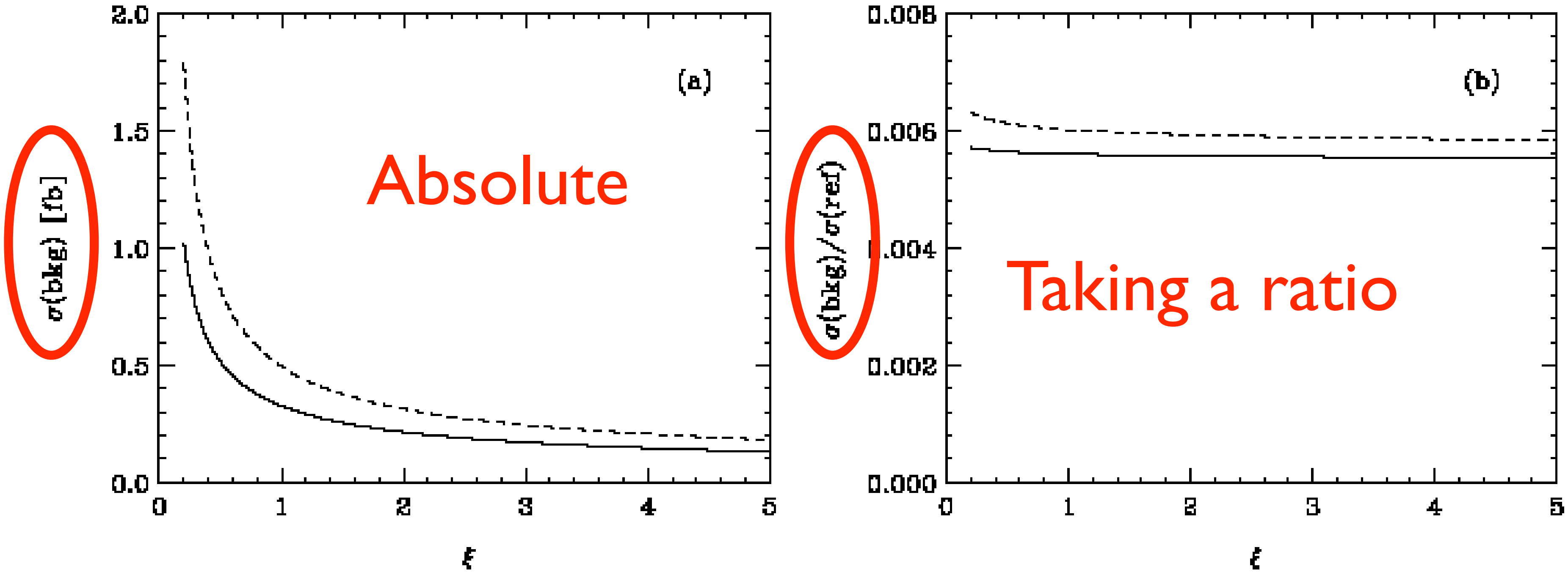


Reducing fake jets



"Alpha parameter" allows to reduce the number of fakes

ttbar theory uncertainties reduced in ratio !



Variation of renormalization and factorization scale

40-50% uncertainty at LO **5%** with the ratio !

[N. Kauer, Phys.Rev.D70, 2004]

Systematic errors on backgrounds ($5fb^{-1}$)

- ww: 17% (dominated by statistics)
- ttbar: 16% (dominated by systematics)
- wz ~20% (could be reduced with further study)
- gg→ww: ~30% (dominated by theoretical errors)
- wtb: ~22% (dominated by theoretical errors)