Estimation of W+jets background using Z+jets as a control sample

Maria Ilaria Besana, Umberto De Sanctis, Tommaso Lari
(Milano group)

CONTENTS:

- Introduction
- Datasets and selection cuts
- W/Z as a function of jet multiplicity
- Differences between rel. 12 and rel. 14
- Estimation of backgroud W+jets to top analysis
- Method to reduce systematic error
- Conclusion and outlook

INTRODUCTION:

- Update of previous analyses:
- http://indico.cern.ch/getFile.py/access?contribId=8&resId=0&materialId=slides&confId=25489
- http://indico.cern.ch/getFile.py/access?contribId=12&resId=0&materialId=slides&confId=25360

- Similar work done by:
- Marisa Sandhoff (Wuppertal) : http://indico.cern.ch/getFile.py/access?contribId=49&sessionId=6&resId=0&materialId=slides&confId=31799
- LBNL group : Atlas note ATL-PHYS-COM-2007-100
- Bonn group :

  http://indico.cern.ch/getFile.py/access?contribId=3&resId=0&materialId=slides&confId=45163

INTRODUCTION: MOTIVATION

- theoretical Xsection of W+jets is rather uncertain especially ratio:

  W+0j : W+1j : W+2j : W+3j : W+4j : W+5j
- tool: Z+jets
  - Z is “easy ” to select with high purity (but $\sigma_W \sim 10 * \sigma_Z$)
  - idea: use Z to determine σ_W and properties of W+jets
  - assumption: $\sigma_{W+4jets} = c * \sigma_{Z+4jets} (\sigma_{W+0jets} / \sigma_{Z+0jets})$

    c = coef which takes into account mass difference for W and Z
- distributions for W and Z should look similar

DATASETS:

Pythia $Z \to ee$	
Z inclusive	106050

Pythia $W \to e\nu$	
W inclusive	06020

Alpgen : $Z \to ee$	
Z + 0 partons	107650
Z + 1 parton	107651
Z + 2 partons	107652
Z + 3 partons	107653
Z + 4 partons	107654
Z + 5 partons	107655

Alpgen : $W \to e\nu$	
W + 0 partons	107680
W + 1 parton	107681
W + 2 partons	107682
W + 3 partons	107683
W + 4 partons	107684
W + 5 partons	107685

Datasets at 10 TeV rel 14: Alpgen datasets haven't any filter on jets.

Selection cuts: object definition

	Reco level		**MC level**
Electrons:	IsEM medium	Electrons:	|pdgId| = 11
	p_T >15 GeV		p_T >15 GeV
	|η|<1.37 U 1.52<|η|<2.5		|η|<1.37 U 1.52<|η|<2.5
	Isolated: etcone20 < 6 GeV		
Jets	Cone4H1TowerJets	Jets	Cone4TruthJets
	Overlap removal: ΔR(jet,e)>0.2		Overlap removal: ΔR(jet,e)>0.2
	p_T >20 GeV		p_T >20 GeV
	|η|<2.5		|η|<2.5
ETmiss	MET_RefFinal	ETmiss	MET_Truth

SELECTION CUTS: EVENT SELECTION

Z→ee:

- Trigger: *e20i*
- 2 electrons with p_T > 20 GeV
- 80 GeV < invariant mass < 100 GeV
- max (|ETmiss$_X$|, |ETmiss$_Y$|) < 3 σ_{gauss}

W→eν:

- Trigger: *e20i*
- 1 electron with p_T > 20 GeV
- ETmiss > 20 GeV

Top selection requires in addition:

- at least 4 jets with p_T >20 GeV & $|\eta|<2.5$
- at least 3 jets with p_T >40 GeV & $|\eta|<2.5$

Selection cuts:

efficiency W selection

Pythia

Cuts	$efficiencies(\%)$	$events\ W \rightarrow e\nu\ at\ 200pb^{-1}$
Tot	87.66	2835360
$Trigger$	48.25	1560600
$1\ good\ electron$	38.06	1231070
$ETmiss > 20GeV$	34.41	1112840

Alpgen

Cuts	$efficiencies$	$events\ W \rightarrow e\nu\ at\ 200pb^{-1}$
Tot	-	3234180
$Trigger$	47.68	1541970
$1\ good\ electron$	37.71	1219710
$ETmiss > 20GeV$	34.23	1106950

Selection cuts:

Efficiency Z selection

Pythia

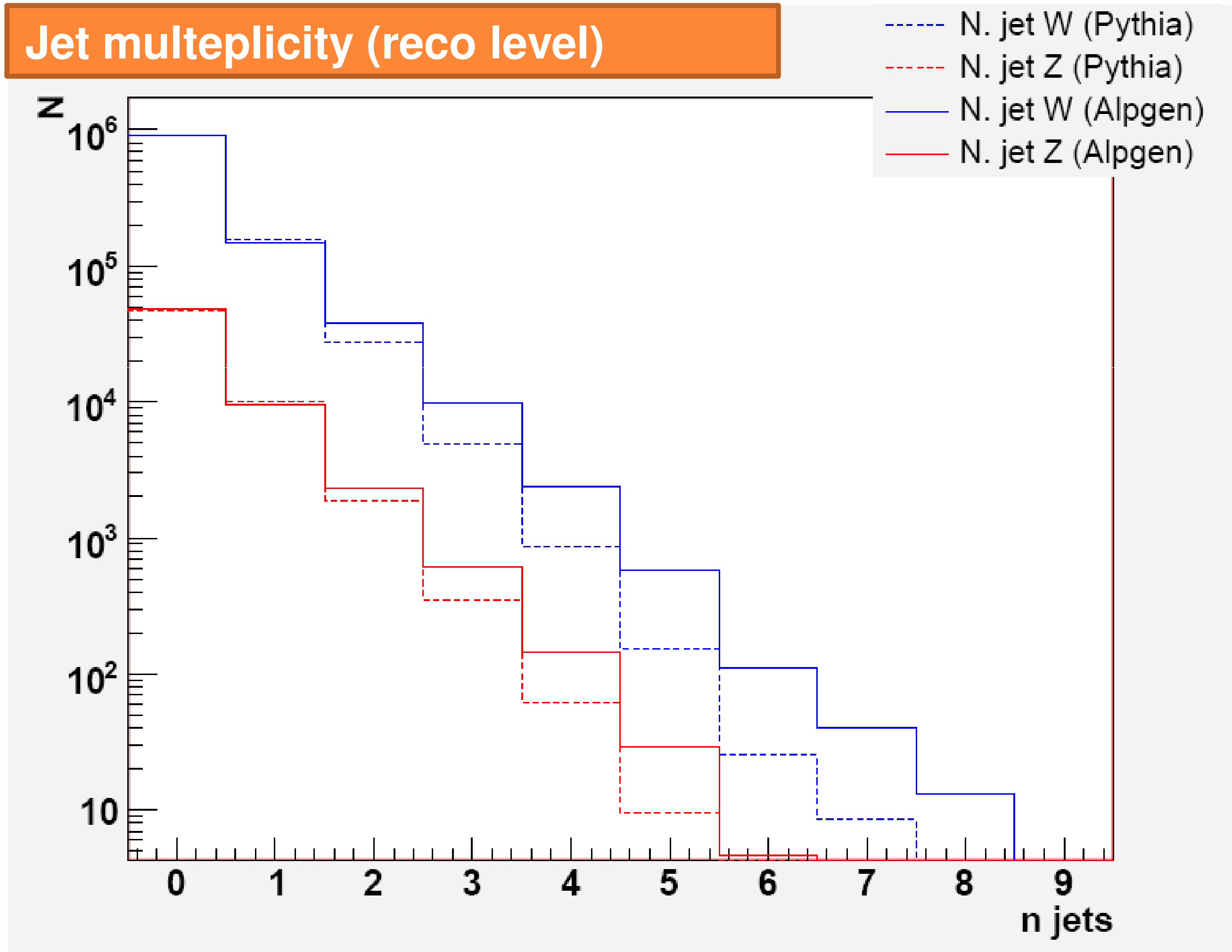
Cuts	$efficiencies(\%)$	$events\ Z \rightarrow ee\ at\ 200pb^{-1}$
Tot	96.01	275704
$Trigger$	67.72	194461
$2\ good\ electrons$	23.46	67380
$80GeV < Lepton\ invariant\ massa < 100GeV$	21.05	60436
$not\ in\ tails\ of\ ETmiss\ distribution$	20.79	59695

Alpgen

Cuts	$efficienze$	$eventi\ Z \rightarrow ee\ attesi\ 200pb^{-1}$
Tot	-	288813
$Trigger$	66.94	193336
$2\ good\ electrons$	23.68	68391
$80GeV < Lepton\ invariant\ massa < 100GeV$	21.33	61618
$not\ in\ tails\ of\ ETmiss\ distribution$	20.99	60613

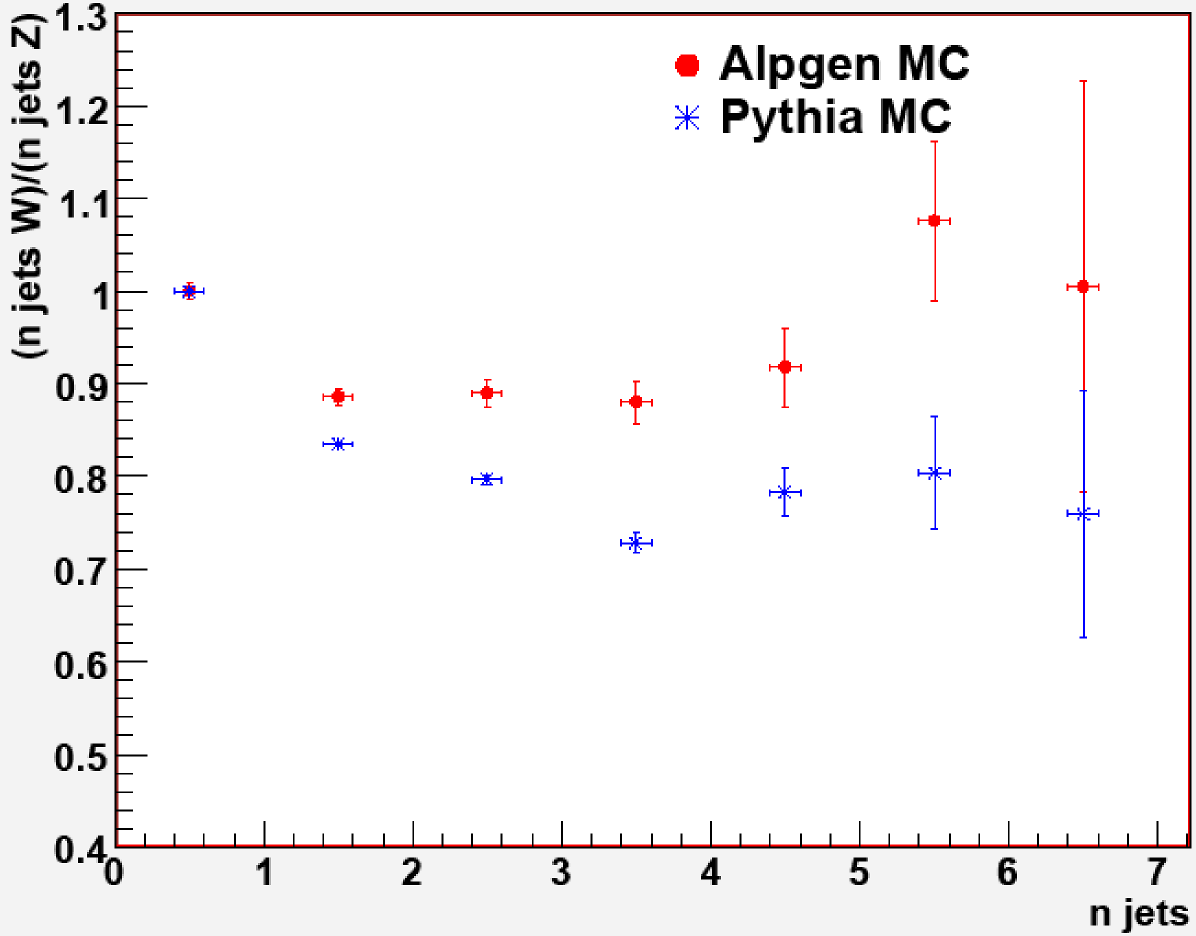
Ratio (signal/background) > 200

JET MULTEPLICITY



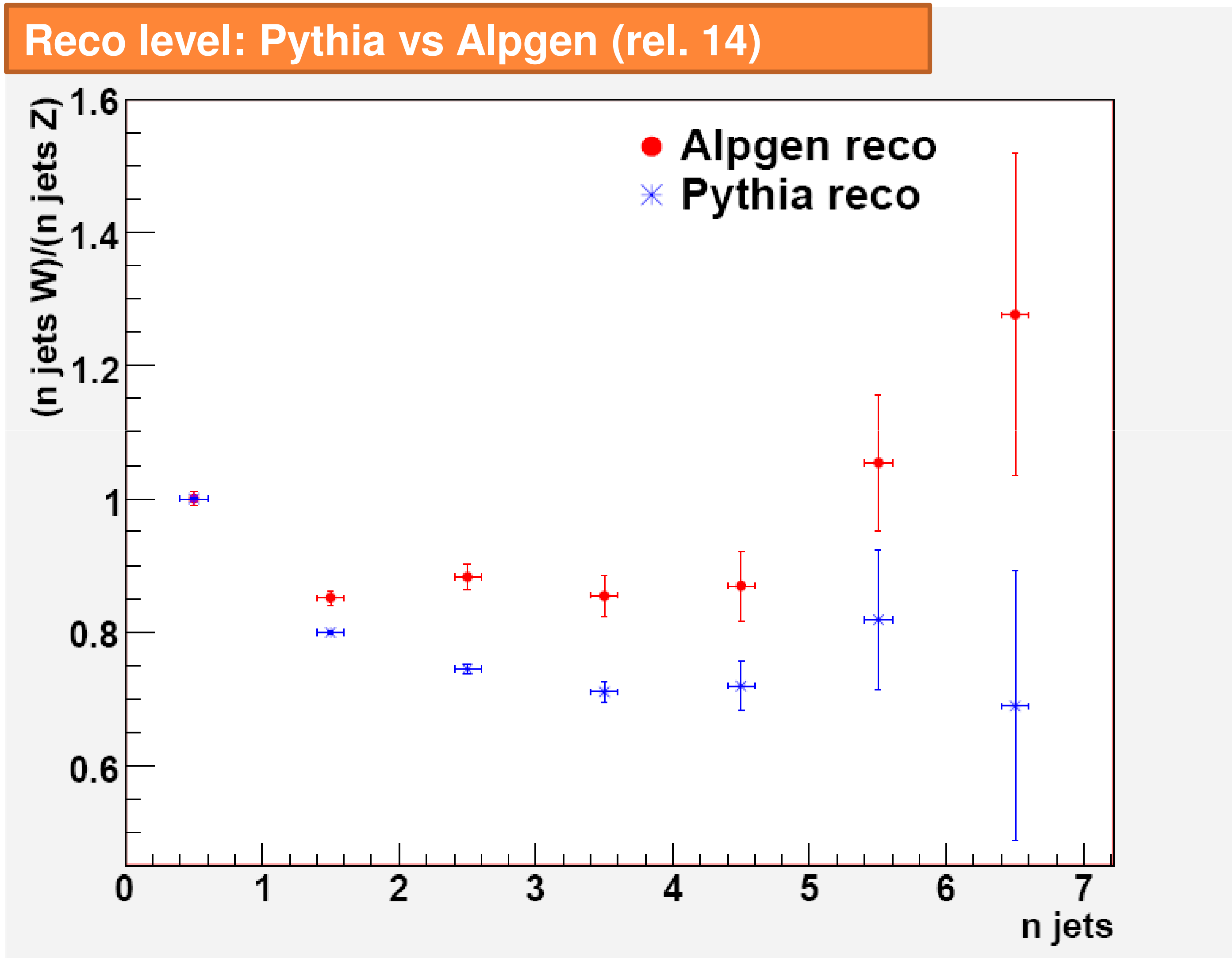
W/Z RATIO AS A FUNCTION OF N_{JETS}:

MC level: Pythia vs Alpgen (rel. 14)



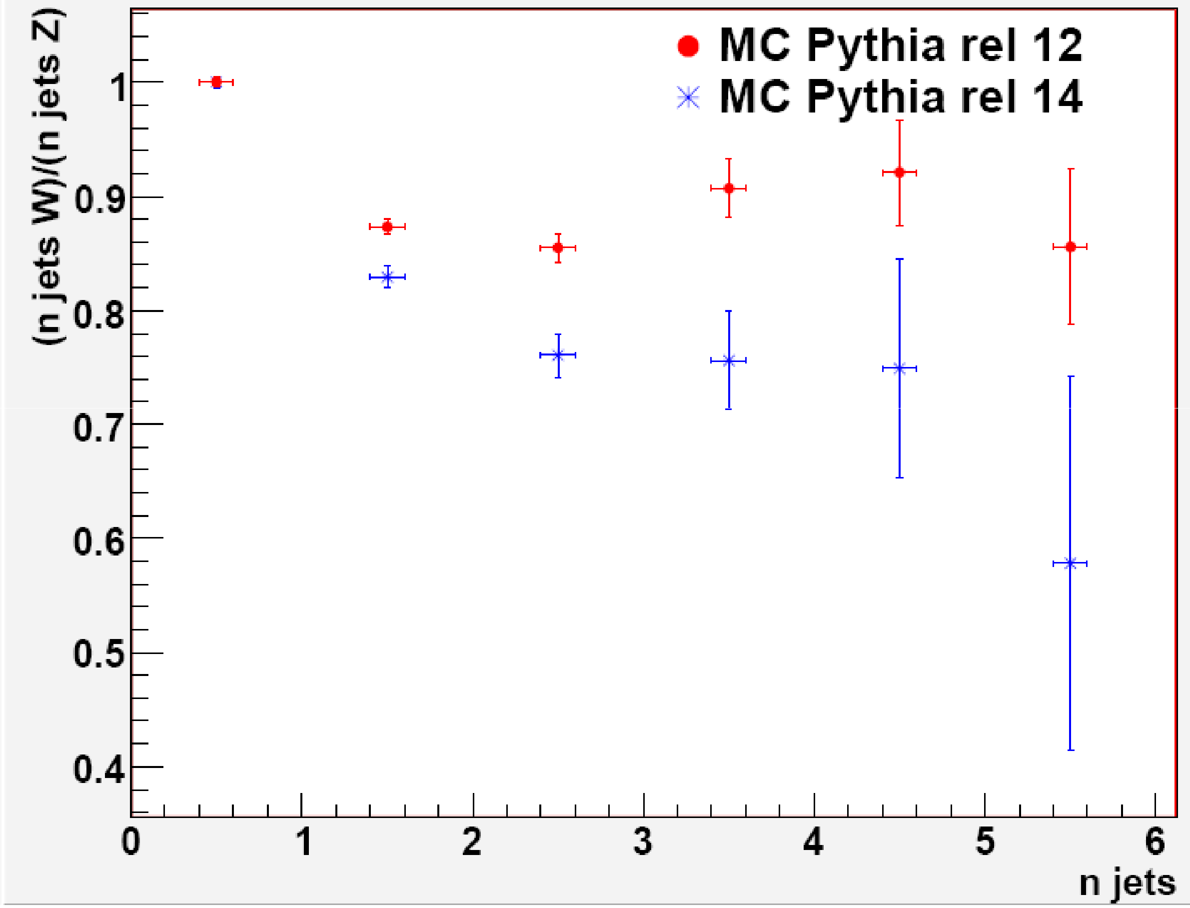
13/03/2009 I. Besana T.Lari U.De Sanctis

W/Z ratio as a function of N_{JETS} :



W/Z RATIO AS A FUNCTION OF N_{JETS} :

MC level: rel. 14 vs rel 12 (Pythia)

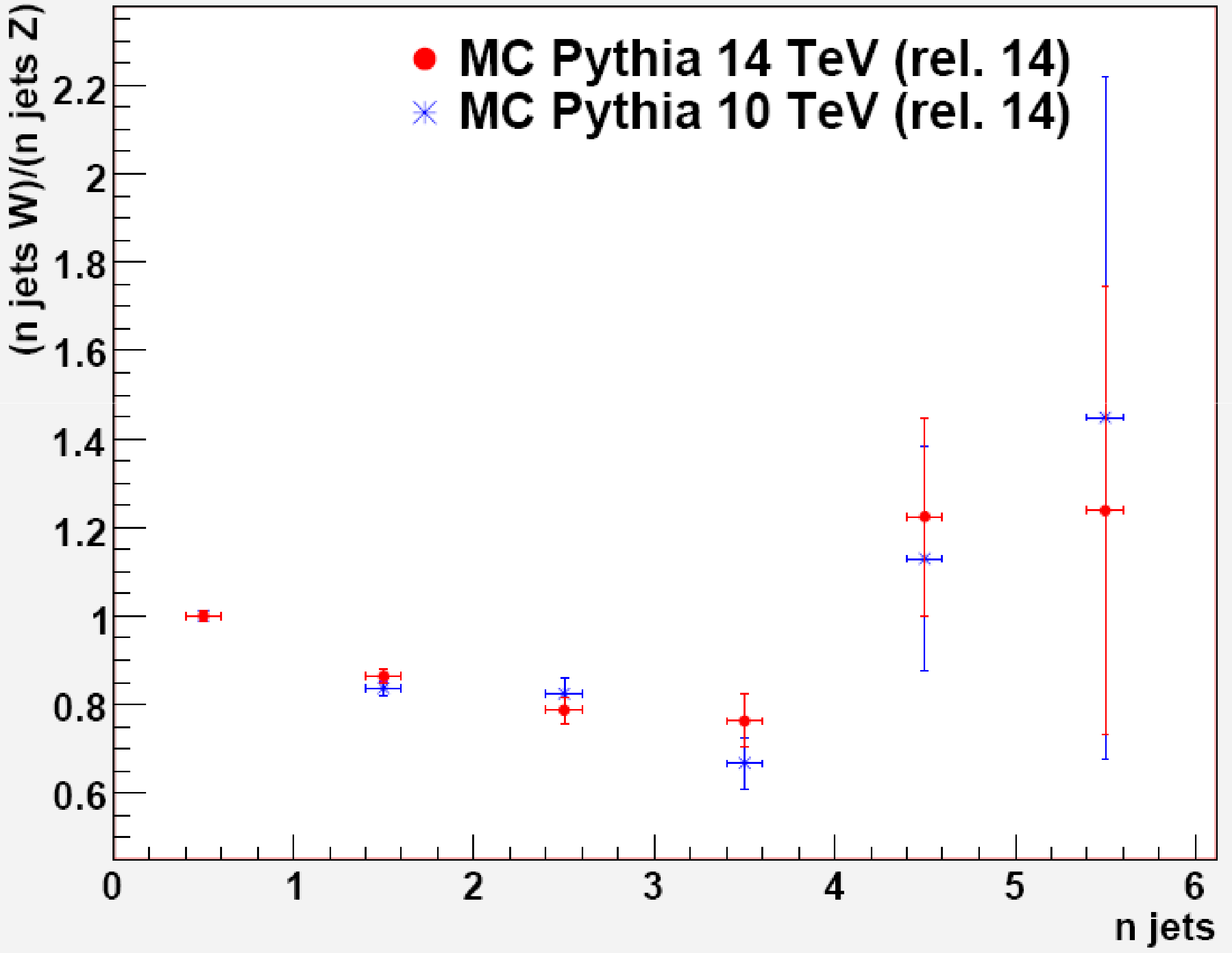


- Lepton trigger: e25
- p_T lepton cut & ETmiss cut: 25 GeV

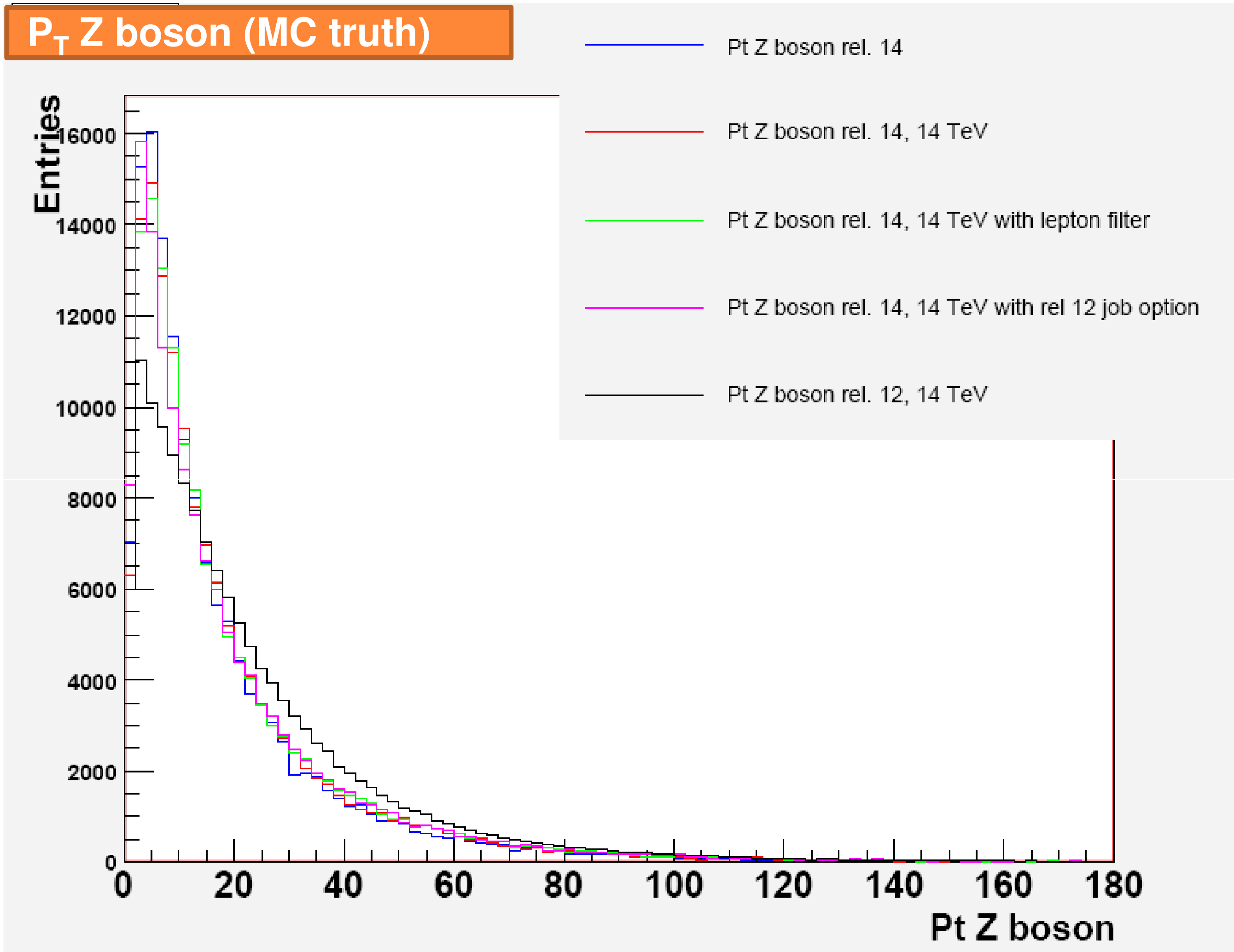
! Not for Alpgen, because Alpgen datasets of release 12 have a filter on jets.

W/Z RATIO AS A FUNCTION OF N_{JETS} :

MC level: rel. 14 TeV vs 10 TeV (Pythia rel 14)

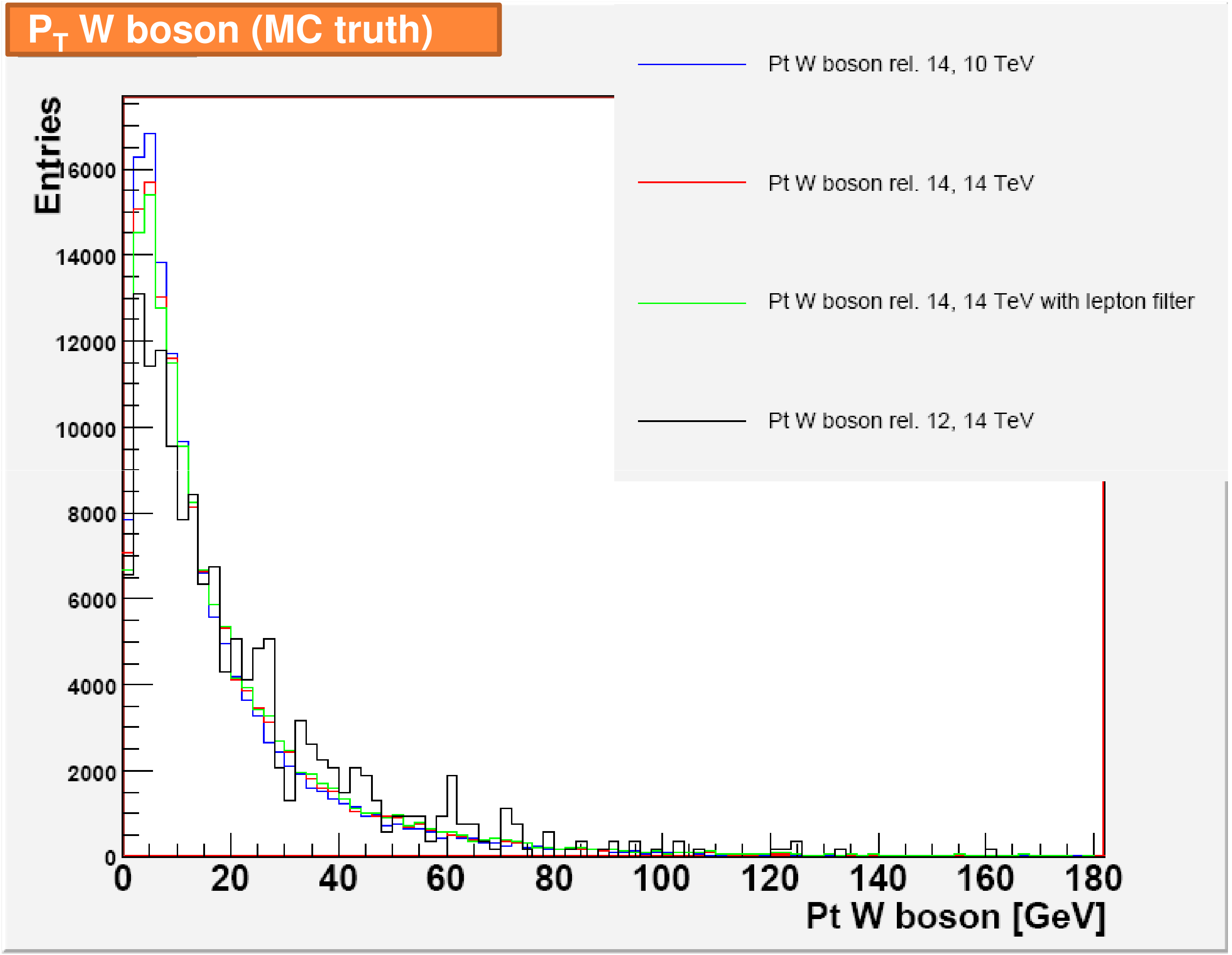


BOSON P_T DISTRIBUTION



Lepton filter: at least 1 lepton with p_T >10 GeV & $|\eta|$< 2.8

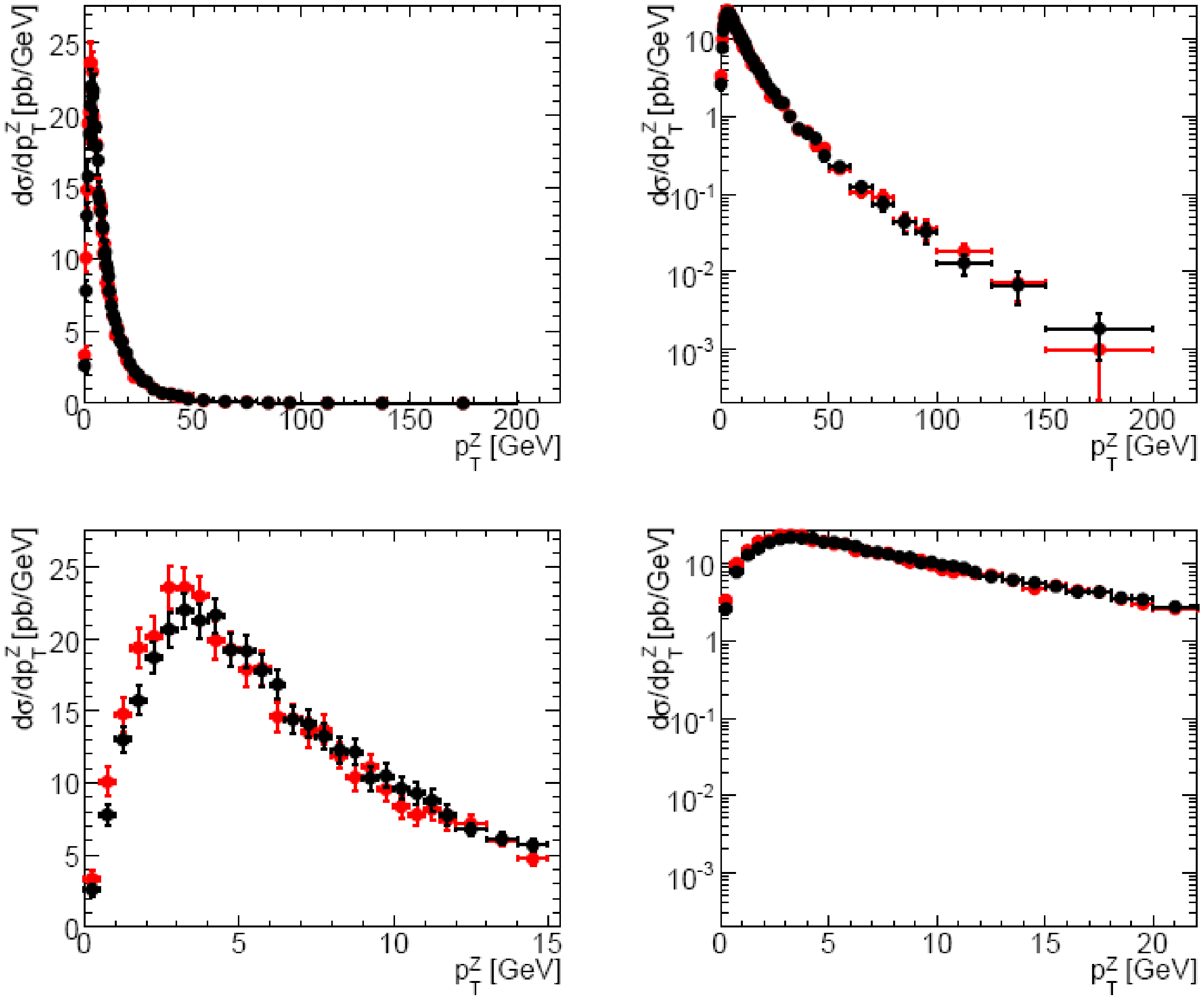
Boson P_T distribution



Lepton filter: at least 1 lepton with p_T >10 GeV & $|\eta|$< 2.8

Boson p_T distribution

Marcus Warsinsky



Red points are the CDF data, black points is Pythia 6.419 with the mc08 tune, but used for Tevatron energies and ppbar.

Evaluation of background to semileptonic top-antitop sample

(200 pb^{-1})

Generator	Pythia	Alpgen
Z events after top selection cuts	20	65
W events after top slection cuts	279	1205
Estimated background	356	1223
Ratio (estimated back.)/(true back.)	1.30	1.04
Statistical error (200 pb^{-1})	0.3	0.13
Statistical error	0.11	0.06

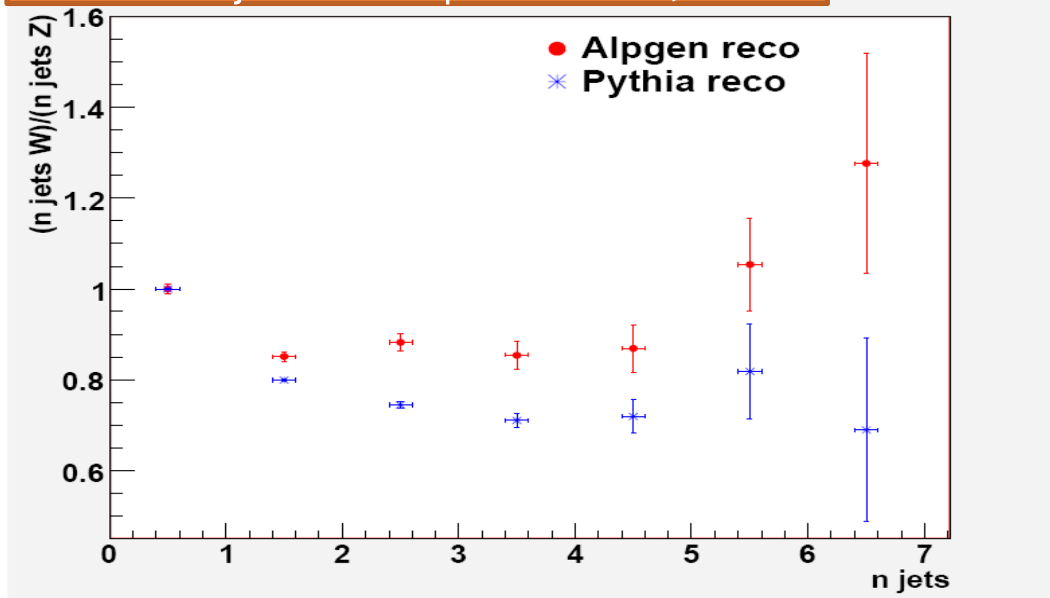
22 % difference between Alpgen and Pyhia

Reduction of systematic error

- Problem: our determination of W+jets background relies on the W/Z ratio as a function of the number of jets being constant or at least known.
- The spread between different generators represents a systematic error, which is at least ~20% for the ttbar signal phase space.
- However, predictions for the ratio already starts to diverge for 1-jet and 2-jet

  (using 0 jet to normalize).
- For this jet multeplicity we should be able to select a relatively pure Z+jets and W+jets sample, with little ttbar contamination.
- The idea is thus to measure R_1/R_0 and R_2/R_0 (R_N=W+Njet/Z+Njet) to constraint the spread of Montecarlo predictions and reduce associated systematics.

Number of jets is with pt > 20 GeV, eta<2.5



Event Selection cuts

- e10i trigger chain (10**31 menu)
- Electrons medium and isolated as before
- W selection: 1 electron pt >20 GeV, no muon pt>20GeV, etmiss > 20 GeV, mT(e,EtMiss) cut
- Z selection: 2 electrons pt>20 GeV, no muon pt>20GeV,
- PxMiss and PyMiss less than 3s from 0, m(ee) cut.

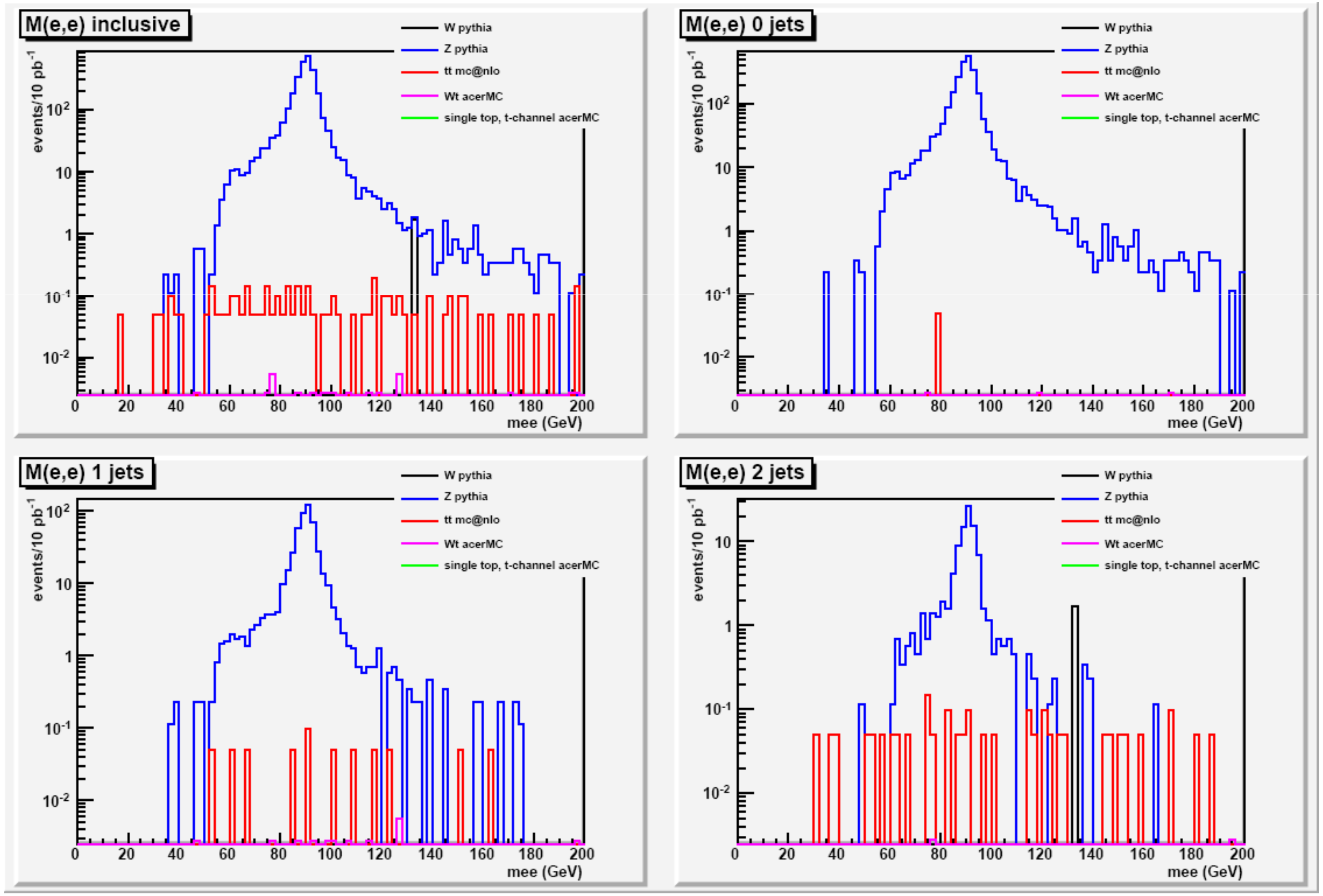
ttbar, W+jets (pythia), Z+jets (pythia) single top considered

Not checked yet:

- QCD contamination in W and Z selected sample, how to estimate and subtract its contribution from data
- other backgrounds, for exemple W $\rightarrow$ e ν and W $\rightarrow$ $\tau\nu$

Z EVENTS:

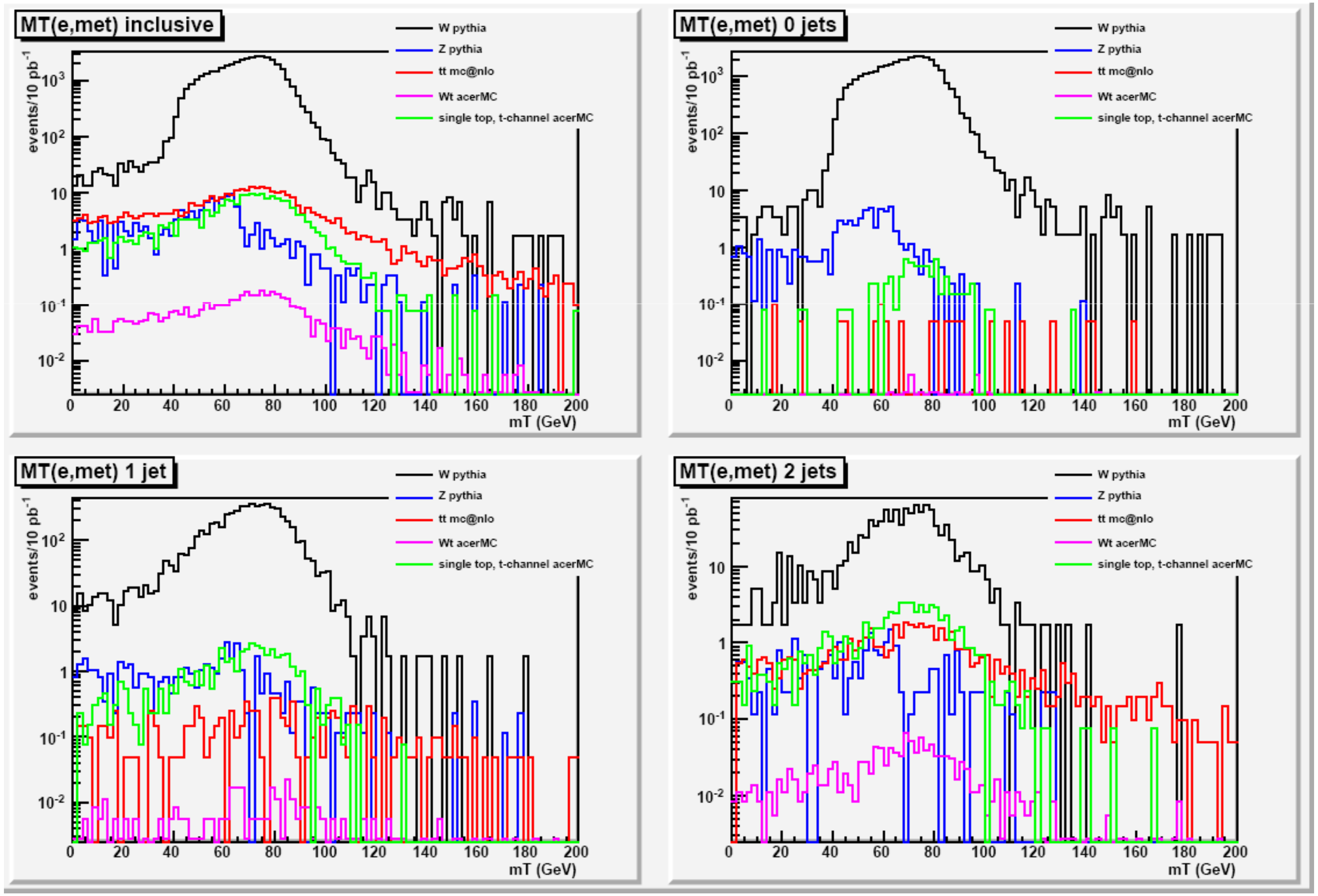
Everything normalized to 10 pb^{-1}
ttbar and single top contamination in Z selection negligible
Selection for ratio measurement: [80 GeV, 100 GeV]



13/03/2009 I. Besana T.Lari U.De Sanctis

W EVENTS:

Everything normalized to 10 pb^{-1}
ttbar and single top contamination in W+2j a few per cent
Selection for ratio measurement: [40 GeV, 90 GeV]



PREDICTIONS FOR R1 AND R2

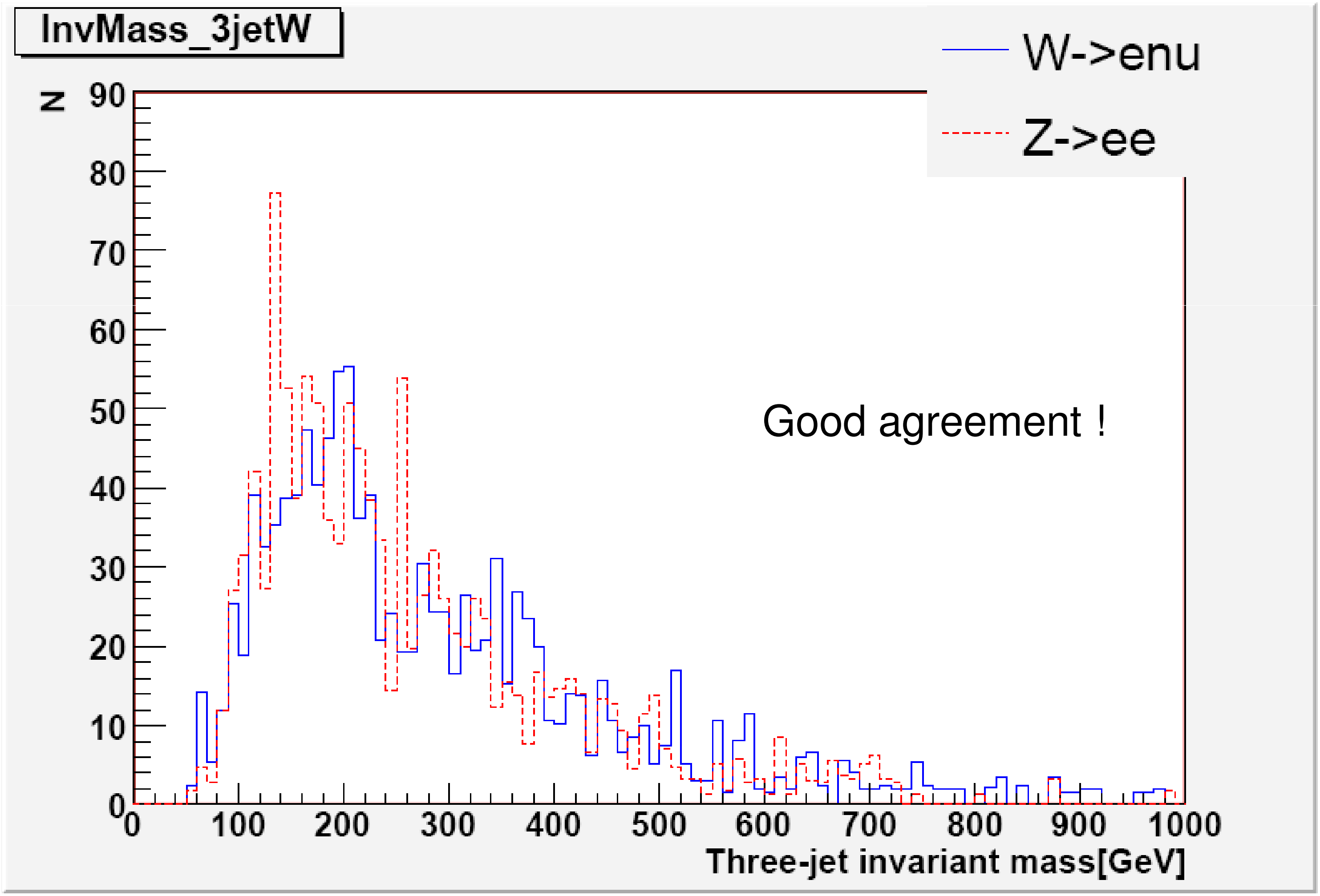
	W,Z events only (with the error from MC statistics)	Effect of backgrounds	Statistical error with 10 pb^{-1}
R1/R0	0.807 ± 0.023	+0.005	0.04
R2/R0	0.76 ± 0.05	+0.046	0.10

The statistical error which can be obtained with 10 pb^{-1} is comparable to the difference between alpgen and pythia predictions, and larger than the effect from ttbar and single top contamination.

It looks promising. With more than 10 pb^{-1}, we may indeed be able to reduce the systematics on our W+jets prediction measuring the W/Z ratio at low jet multeplicity.

THREE-JET INVARIANT MASS DISTRIBUTION (200 PB^{-1})

Alpgen



Conclusion & Outlook

W+jets background can be estimated from Z+jets:

- with a statistical error of 13% (according to Alpgen) at 200 pb^{-1} ,
- difference between Alpgen and Pythia of 22%,
- using R_1/R_0 and R_2/R_0 (R_N=W+Njet/Z+Njet) to constraint the spread of Montecarlo predictions and reduce associated systematics looks promising.

Next steps:

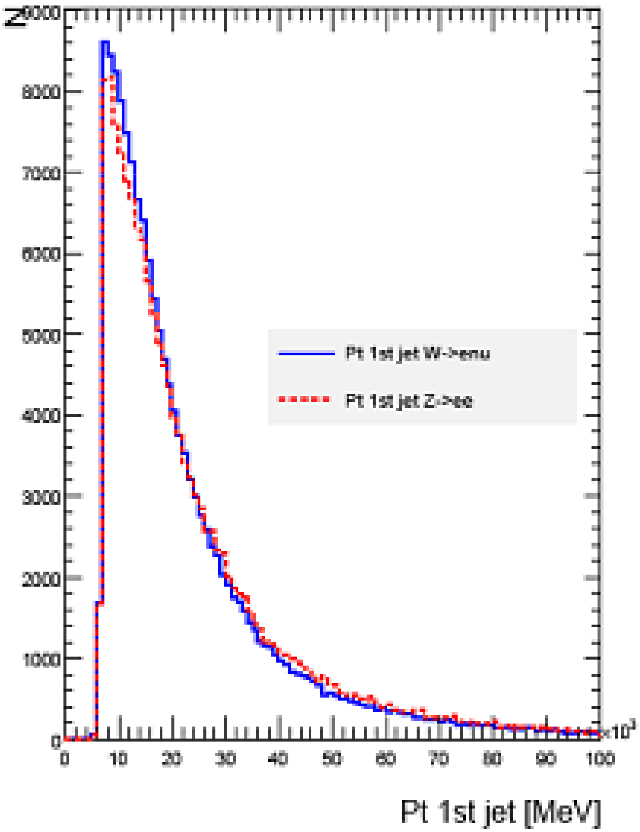
- include background to Z selection,
- study systematics:

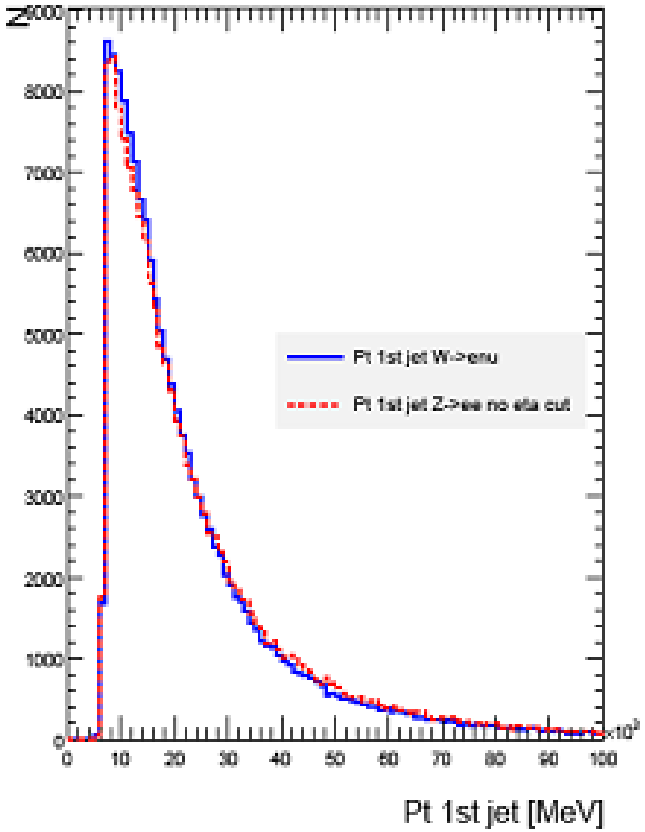
  constrain generator dependence with first data.

Backup

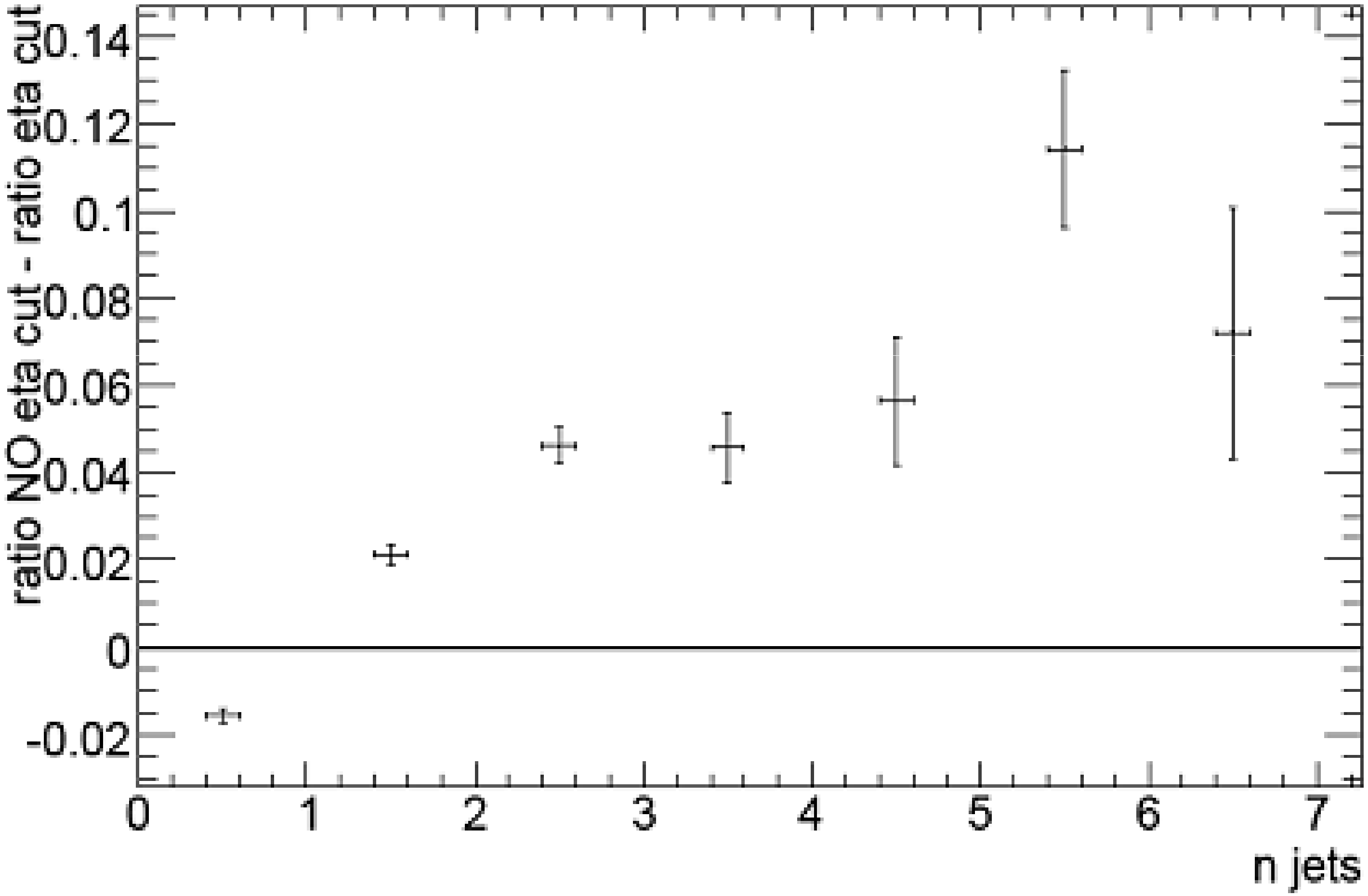
PYTHIA SYSTEMATIC ERROR AT MC LEVEL (REL 12)

There is a distortion in the p_T shape of the most energetic jet wich comes from doing an η cut on the 2nd electron in $Z \longrightarrow ee$, but not on the ν of $W \longrightarrow e\nu$

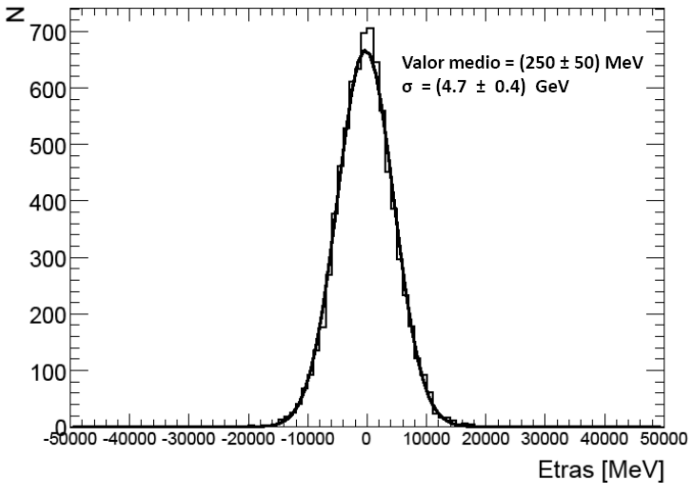




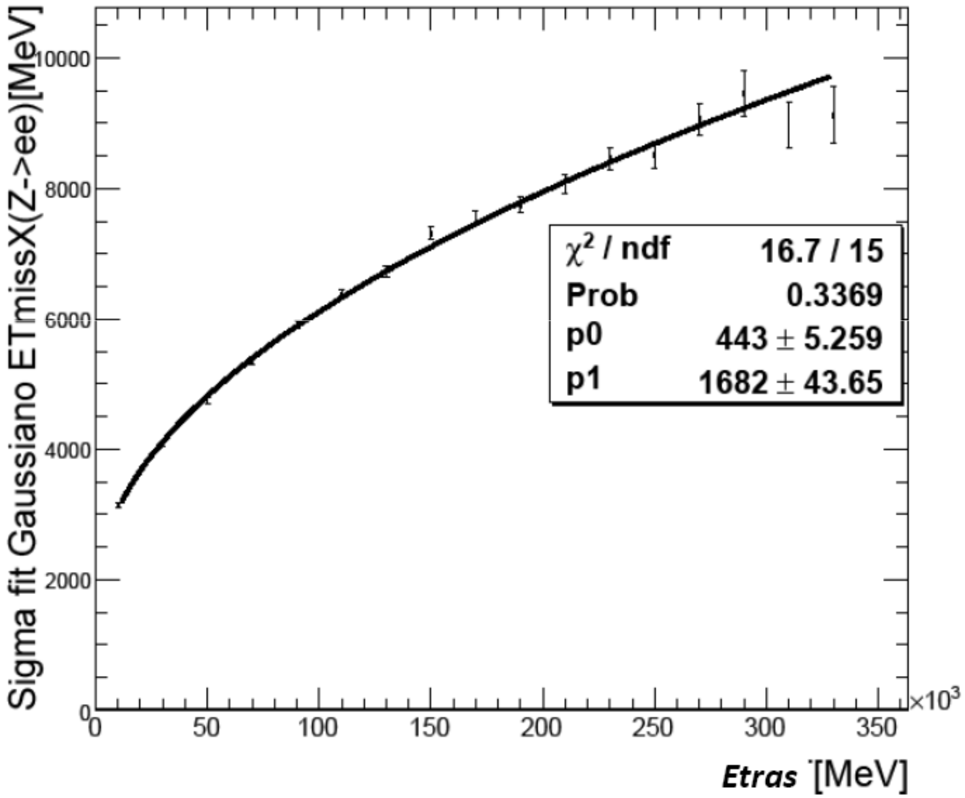
PYTHIA SYSTEMATIC ERROR AT MC LEVEL (REL 12)



ETMISS DISTRIBUTION:



$$\sigma_{gauss} = P_0\sqrt{SumET[GeV]} + P_1$$



Selection cut: max ($|ETmiss_X|$, $|ETmiss_Y|$) $< 3\ \sigma_{gauss}$