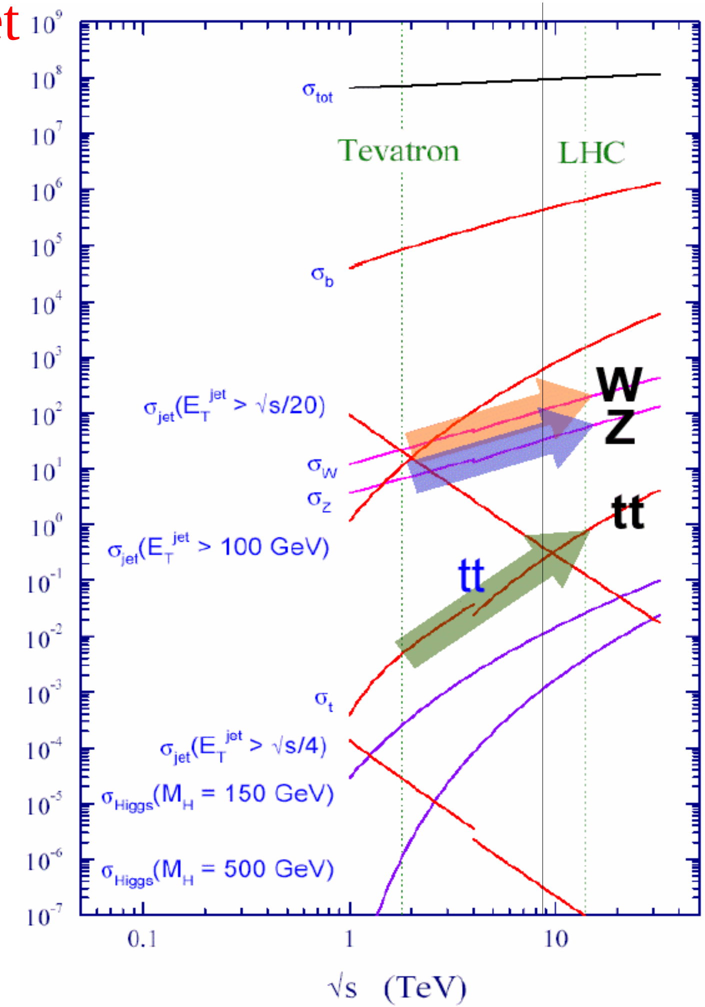
About TAUOLA...

- Ancestors information (parton-level) appears to be missing in the samples where Tauola was used to decay tau's.
- The problem have been located: due to multiple translations between PYJETS and HEPEVT common blocks, via calls to PYHEPC(1) or PYHEPC(2), some records were overridden, and thus correupted mother-daugther links in the hard event block.
- NOTE-1: multiple calls to PYHEPC(1) and PYHEPC(2) were necessary because Tauola operates on HEPEVT standard, but can use Pythia6 to decay remaining unstable hadrons.
- NOTE-2: final state particles NOT affected.
- The problem has been solved by placing a local copy of updated PYHEPC routine (by Steve M.) into GeneratorInterface/CommonInterface package. Tag is ready for publishing, tested with Pythia6+Tauola and MG+Tauola, under 2_1_12 and 2_2_0_pre1.

Multi-jets background estimation from data

Aim : to estimate, from data, the number of QCD di-jet events in the context of semi-muonic ttbar analysis with a not fully understood detector.



Outlines :

- Technicalities
- Standard ABCD method :
  - Principle
  - Aim and observables used
  - Results
- Conclusions
- More if we have times...

Technicalities

- CMSSW 1_6_10, PAT_1610_080229
- Background and signal sample :
  - QCD PYTHIA "/ppMuPt20-15/CMSSW_1_6_7-CSA07-1204357236/RECO" :
    - Filtered at generator level for a muon with p_T>15GeV/c and p_T_hat >20GeV/c
    - ~2M events = 8.7pb^{-1} (rescaled to 10pb^{-1})
    - 100 pb^{-1} misalignment, miscalibration scenario
  - Chowder alpgen : tt+jets and W/Z+jets
    - ~1M events (rescaled to 10pb^{-1})
- Event pre-selection :
  - At least one reconstructed muon with p_T>20 GeV/c, |eta|<2.1 (trigger)
  - At least four reconstructed jets with E_T>30 GeV, |eta|<2.4
  - ΔR>0.3 between the selected muons and its closest jets (among the four leading ones)

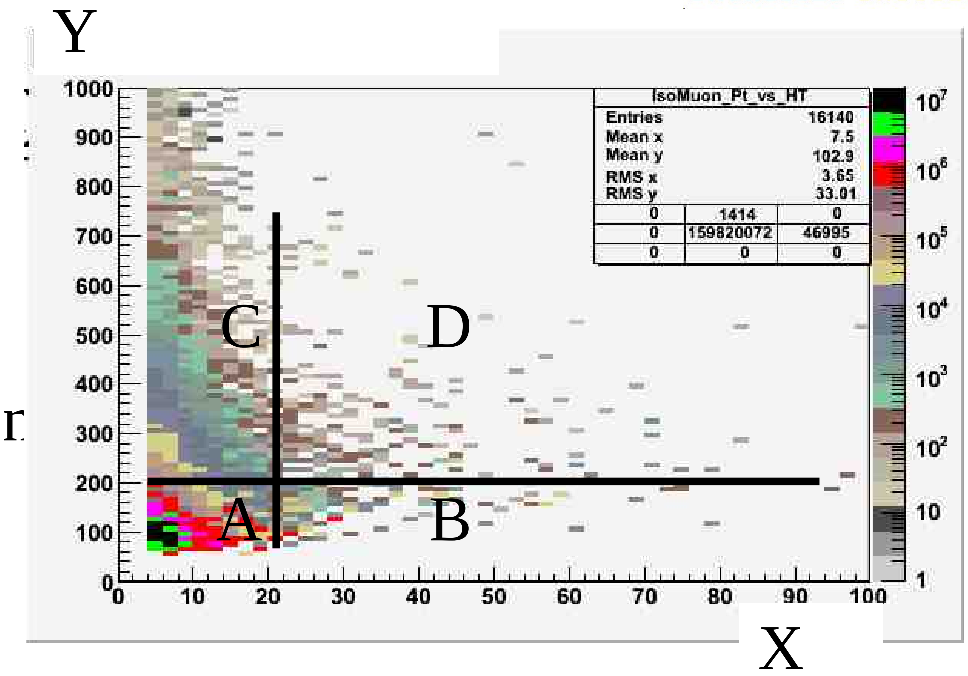
ABCD Method

- Principle of the method :
  - If X and Y are uncorrelated :

    $N_C^{bckgd}/N_D^{bckgd} = N_A^{bckgd}/N_B^{bckgd}$

  - E.g. : D is the signal region, then estimate N_D^{bckgd} :

    $N_D^{bckgd} = N_B^{bckgd} * N_C^{bckgd}/N_A^{bckgd}$



If the regions A, B and C are background dominated, then :

$N_B^{bckgd} \sim N_B^{exp}$, number of events experimentally observed.

Background Estimation from data

- Aim :
  - Select semi-mu tt+jets events with first data : detector not well understood. Avoid use of missing E_T, b-tagging,...
- Observables investigated :

Lepton related variables :

- p_T of the leading muon
- Calorimeter isolation
- Tracker isolation
- Combined isolation
- RelIso (p_T/(Calo+TrackIso+p_T))

Jet related variables :

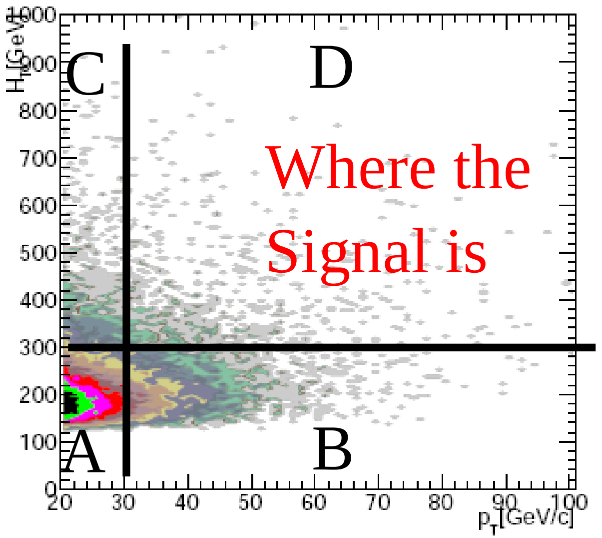
- H_T (scalar sum of the E_T of all the jets E_T>30 GeV)
- $E_T^{3,4}$ (scalar sum of the E_T of the 3rd and 4th jet / H_T)

Background Estimation from data

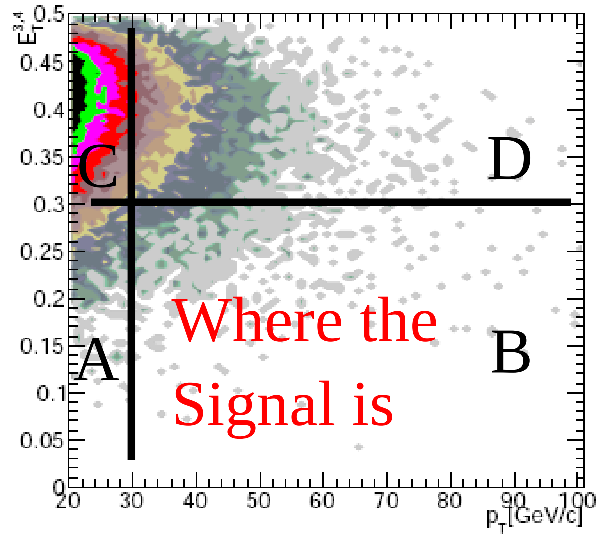
•

H_T vs $p_T(\mu)$	$E_T^{3,4}$ vs $p_T(\mu)$

QCD



QCD



	$p_T^\mu < 30\ GeV/c$	$p_T^\mu > 30\ GeV/c$
$H_T > 300\ GeV$	49 ± 7	17 ± 4 Estimation : 13 ± 2
$H_T < 300\ GeV$	788 ± 28	206 ± 14

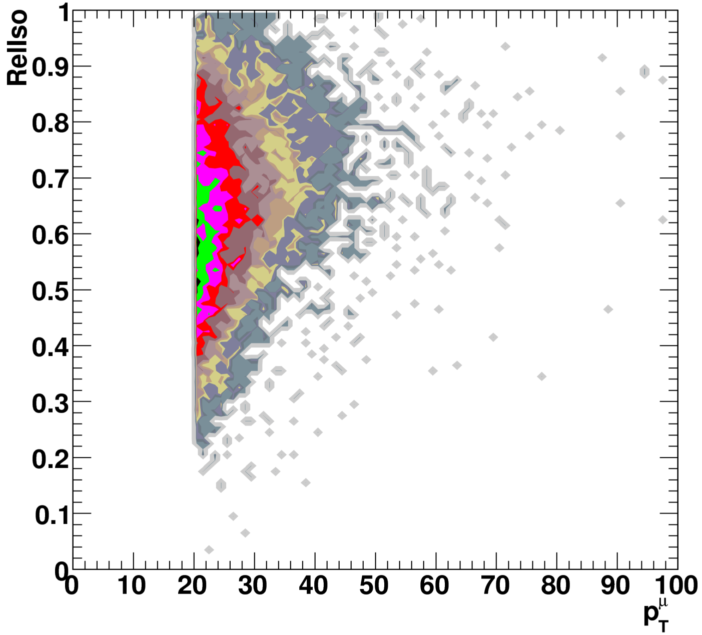
	$p_T^\mu < 30\ GeV/c$	$p_T^\mu > 30\ GeV/c$
$E_T^{3,4} > 0.35$	564 ± 24	140 ± 12
$E_T^{3,4} < 0.35$	274 ± 17	83 ± 9 Estimation : 68 ± 8

Too strong correlation between muon and jet transverse energies :

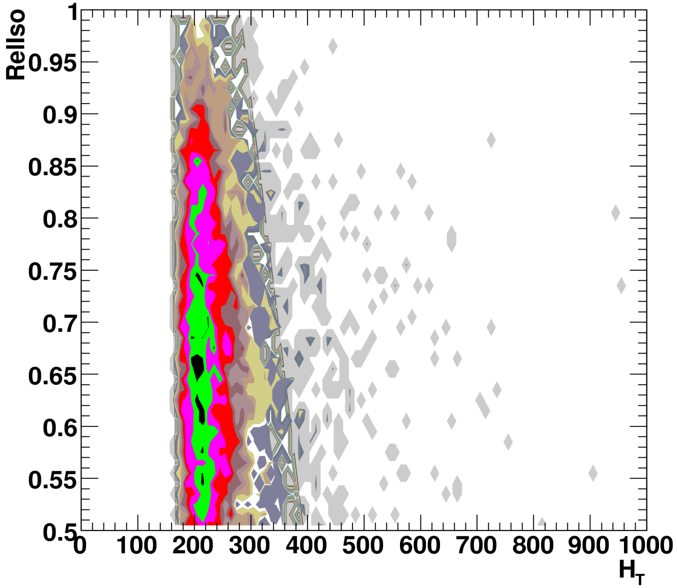
results are not reliable as they depend on the cuts...

Background Estimation from data

- RelIso vs $p_T(\mu)$ QCD



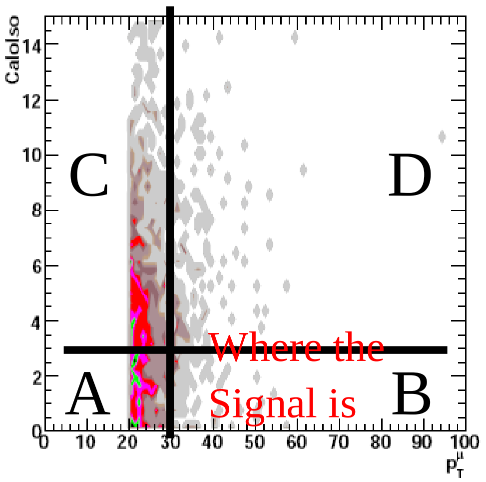
RelIso vs H_T QCD



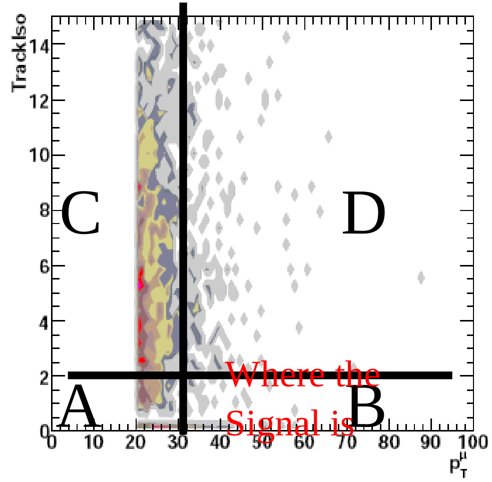
RelIso has a nice rejection power but does not suit for background estimate with the ABCD method...

Background Estimation from data

CaloIso vs...



TrackIso vs ...



	$p_T^\mu < 30\ GeV/c$	$p_T^\mu > 30\ GeV/c$
$CaloIso > 3$	635 ± 25	185 ± 13
$CaloIso < 3$	383 ± 20	92 ± 10 Estimation : 112 ± 11
$TrackIso > 2$	1405 ± 38	340 ± 18
$TrackIso < 2$	383 ± 20	92 ± 10 Estimation : 93 ± 7

- tt+jets : 290(±17)

- W/Z+jets : 115(±11)/12(±3)

Isolation variables seem to be less correlated to the muon transverse momentum...

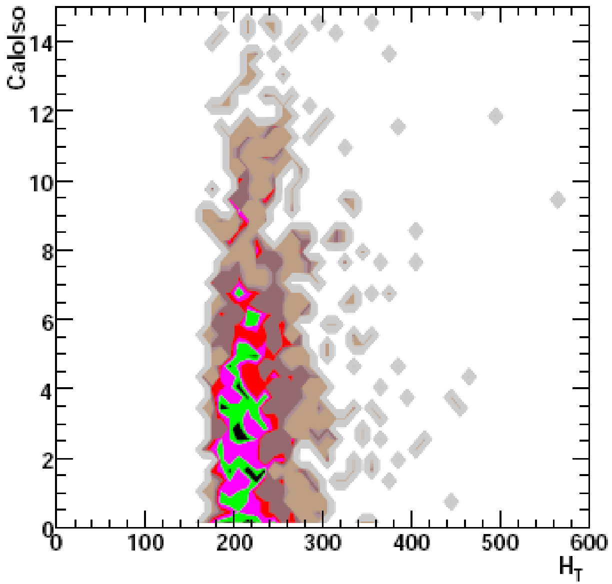
Reasonable agreement with the expected number of bckgd events, $S_{tt+jets}/N_{QCD} = 3.2$

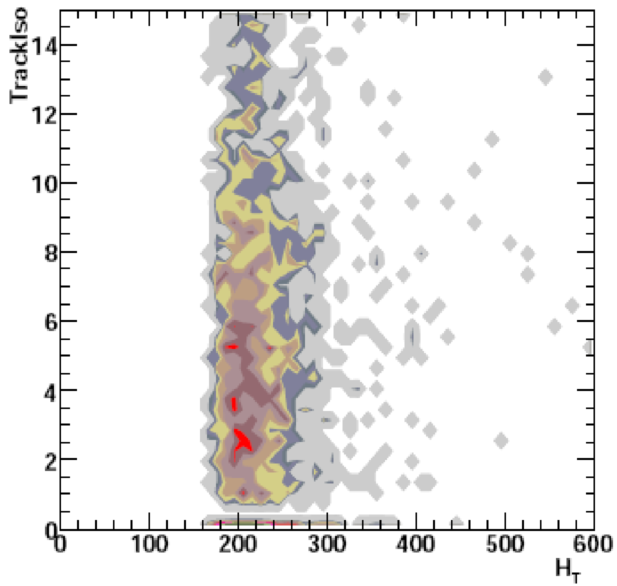
But...

p_T>30 & CaloIso<2 : Est 53±7 / Exp 79±8

p_T>30 & CaloIso<4 : Est 128±11 / Exp 151±16

Background Estimation from data





TrackIso seems less correlated with H_T or $E_T^{3,4}$ than CaloIso...

- tt+jets : 210±15
- W/Z+jets :74±9/7±3
- $S_{tt+jets}/N_{QCD}$ = 6.6

	$H_T < 300$ GeV	$H_T > 300$ GeV	$E_T^{3,4} < 0.35$	$E_T^{3,4} > 0.35$
CaloIso > 3 GeV	743 ± 27	77 ± 9	263 ± 16	557 ± 24
CaloIso < 3 GeV	443 ± 21	32 ± 6 Estimation : 46 ± 6	145 ± 12 Estimation : 156 ± 14	330 ± 18
TrackIso > 2 GeV/c	1634 ± 40	112 ± 11	620 ± 25	1126 ± 34
TrackIso < 2 GeV/c	443 ± 21	32 ± 6 Estimation : 30 ± 3	145 ± 12 Estimation : 182 ± 14	330 ± 18

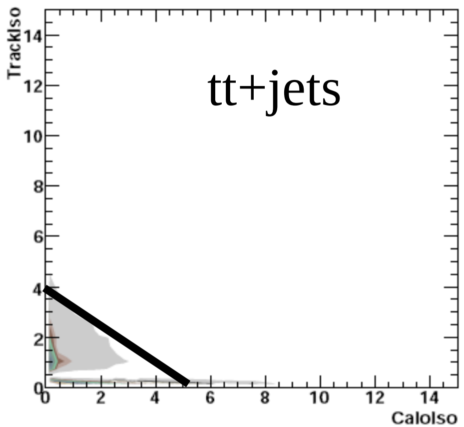
TrackIso vs H_T seems to produce an accurate and stable estimate :

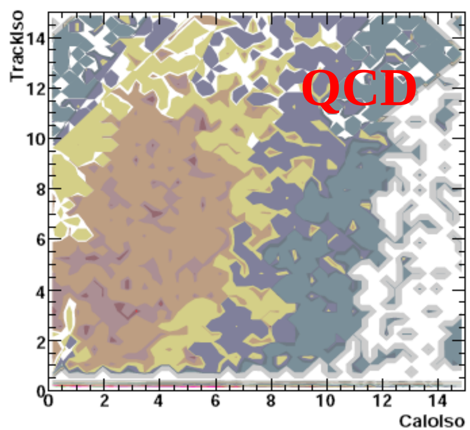
TrackIso<3 : Est 41±6 Exp 44±4

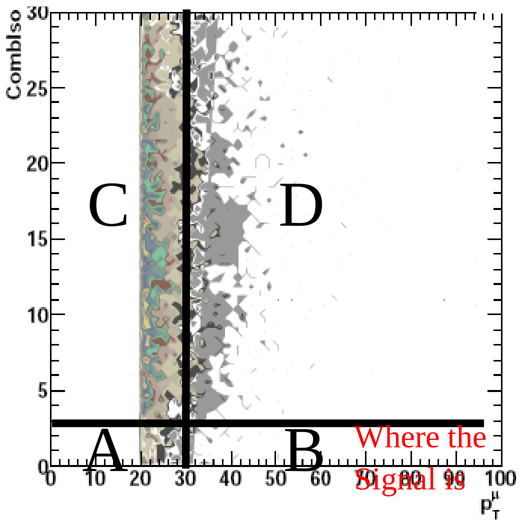
TrackIso<1 : Est 17±4/ Exp 22±4

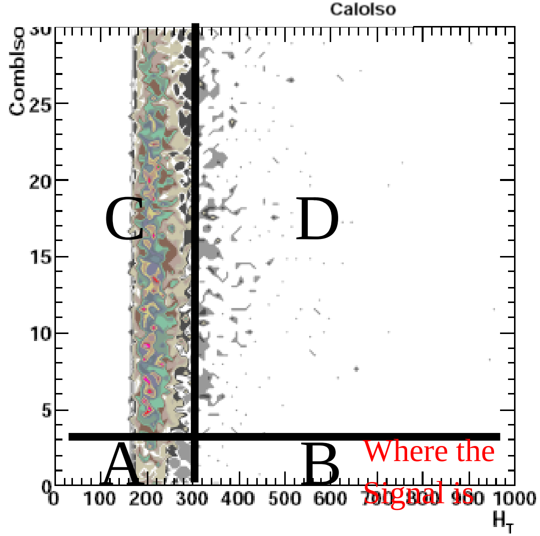
Background Estimation from data

- Combined isolation : CombIso = ((3/2) TrackIso + CaloIso)









	$p_T^\mu < 30\ GeV/c$	$p_T^\mu > 30\ GeV/c$
$CombIso > 3$	6554 ± 81	2019 ± 45
$CombIso < 3$	277 ± 17	72 ± 9 Estimation : 85 ± 6

	$H_T < 300\ GeV$	$H_T > 300\ GeV$
$CombIso > 3$	7527 ± 87	1047 ± 32
$CombIso < 3$	332 ± 18	17 ± 4 Estimation : 46 ± 3

- Still in good agreement with the muon p_T
  - Stable against variations of the cuts
  - Est/Exp : 24±5/25±3 if CombIso<1
  - $S_{tt+jets}/N_{QCD}$: 220/24~9.2
  - But :
    - Is the estimate robust against the signal spillage in bckgd regions?

Background Estimation from data

- Signal spillage :
  - QCD :

	$p_T^\mu < 30\ GeV/c$	$p_T^\mu > 30\ GeV/c$
$CombIso > 3$	6554 ± 81	2019 ± 45
$CombIso < 3$	277 ± 17	72 ± 9 Estimation : 85 ± 6

	$H_T < 300\ GeV$	$H_T > 300\ GeV$
$TrackIso > 2\ GeV/c$	1634 ± 40	112 ± 11
$TrackIso < 2\ GeV/c$	443 ± 21	32 ± 6 Estimation : 30 ± 3

  - QCD + « tt+jets » + « W/Z+jets »

$t\bar{t} + jets$, $W/Z + jets$ and QCD	$p_T^\mu < 30\ GeV/c$	$p_T^\mu > 30\ GeV/c$
$CombIso > 3$	6591 ± 81	2217 ± 46
$CombIso < 3$	368 ± 19	$t\bar{t}$: 281 ± 17, $W/Z + jets$: $111 \pm 11/12 \pm 3$ Expected QCD : 72 ± 9 Estimated QCD : 118 ± 7

$t\bar{t} + jets$, $W/Z + jets$ and QCD	$H_T < 300\ GeV/c$	$H_T > 300\ GeV/c$
$TrackIso > 2$	1651 ± 41	136 ± 12
$TrackIso < 2$	663 ± 26	$t\bar{t}$: 210 ± 15, $W/Z + jets$: $74 \pm 9/7 \pm 3$ Expected QCD : 32 ± 6 Estimated QCD : 55 ± 5

contribution from tt+jets and W/Z+jets not negligible anymore...

Background Estimation from data

- Idea : separate further signal and background regions...
  - H_T vs TrackIso :

$t\bar{t}$ + *jets*, *W*/*Z* + *jets* and QCD	$H_T < 300\ GeV/c$	$H_T > 300\ GeV/c$
$TrackIso > 2$	1651 ± 41	136 ± 12
$TrackIso < 2$	663 ± 26 $t\bar{t}$: 145 ± 12	$t\bar{t}$: 210 ± 15, $W/Z + jets$: $74 \pm 9/7 \pm 3$ Expected QCD : 32 ± 6 Estimated QCD : 55 ± 5

$t\bar{t}$ + *jets*, *W*/*Z* + *jets* and QCD	$H_T < 200\ GeV/c$	$H_T > 300\ GeV/c$
$TrackIso > 2$	484 ± 22	136 ± 12
$TrackIso < 2$	151 ± 12 $t\bar{t}$: 6 ± 2	$t\bar{t}$: 210 ± 15, $W/Z + jets$: $74 \pm 9/7 \pm 3$ Expected QCD : 32 ± 6 Estimated QCD : 42 ± 5

Helps in reducing the signal spillage in the background region...

#jets ==4 : Exp/Est = 21±5/22±3 // #jets ==5 : Exp/Est = 9±3/17±5

#jets >=6 : Exp/Est = 2±2/6±4

- Idea : separate further signal and background regions...
  - P_T vs CombIso :

$t\bar{t}$ + *jets*, *W*/*Z* + *jets* and QCD	$p_T^\mu < 30\ GeV/c$	$p_T^\mu > 30\ GeV/c$
$CombIso > 1$	5369 ± 73	1694 ± 41
$CombIso < 1$	158 ± 13 $t\bar{t}$: 50 ± 7	$t\bar{t}$: 220 ± 15, $W/Z + jets$: $87 \pm 9/9 \pm 3$ Expected QCD : 24 ± 5 Estimated QCD : 50 ± 4

$t\bar{t}$ + *jets*, *W*/*Z* + *jets* and QCD	$p_T^\mu < 25\ GeV/c$	$p_T^\mu > 30\ GeV/c$
$CombIso > 1$	3742 ± 61	1694 ± 41
$CombIso < 1$	102 ± 10 $t\bar{t}$: 25 ± 5	$t\bar{t}$: 220 ± 15, $W/Z + jets$: $87 \pm 9/9 \pm 3$ Expected QCD : 24 ± 5 Estimated QCD : 46 ± 5

	$p_T^\mu < 30\ GeV/c$	$p_T^\mu > 40\ GeV/c$
$CombIso > 1$	5370 ± 73	486 ± 22
$CombIso < 1$	158 ± 13 $t\bar{t}$: 50 ± 7	$t\bar{t}$: 173 ± 13, $W/Z + jets$: $70 \pm 8/7 \pm 3$ Expected QCD : 8 ± 3 Estimated QCD : 14 ± 2

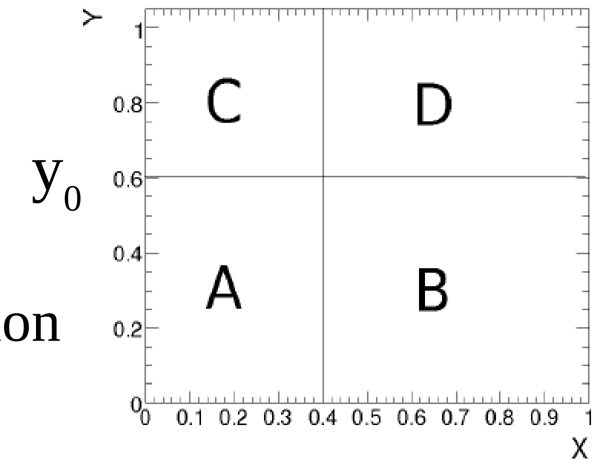
Conclusions...

- ABCD method :
  - p_T^{μ} : too correlated to H_T or $E_T^{3,4}$, performs better with isolation variables and H_T or $E_T^{3,4}$: only TrackIso produced a correct and stable estimation
    - Both suffer from signal spillage in background dominated regions...
  - In fact, the methods works as long as signal is not included...
  - Adding « no man's land » between signal and background regions helps in reducing the signal spillage...
  - What is impact of 14TeV -> 10TeV ?
  - Need to be repeated with 2XX versions...

Back-up slides....

Background Estimation from data

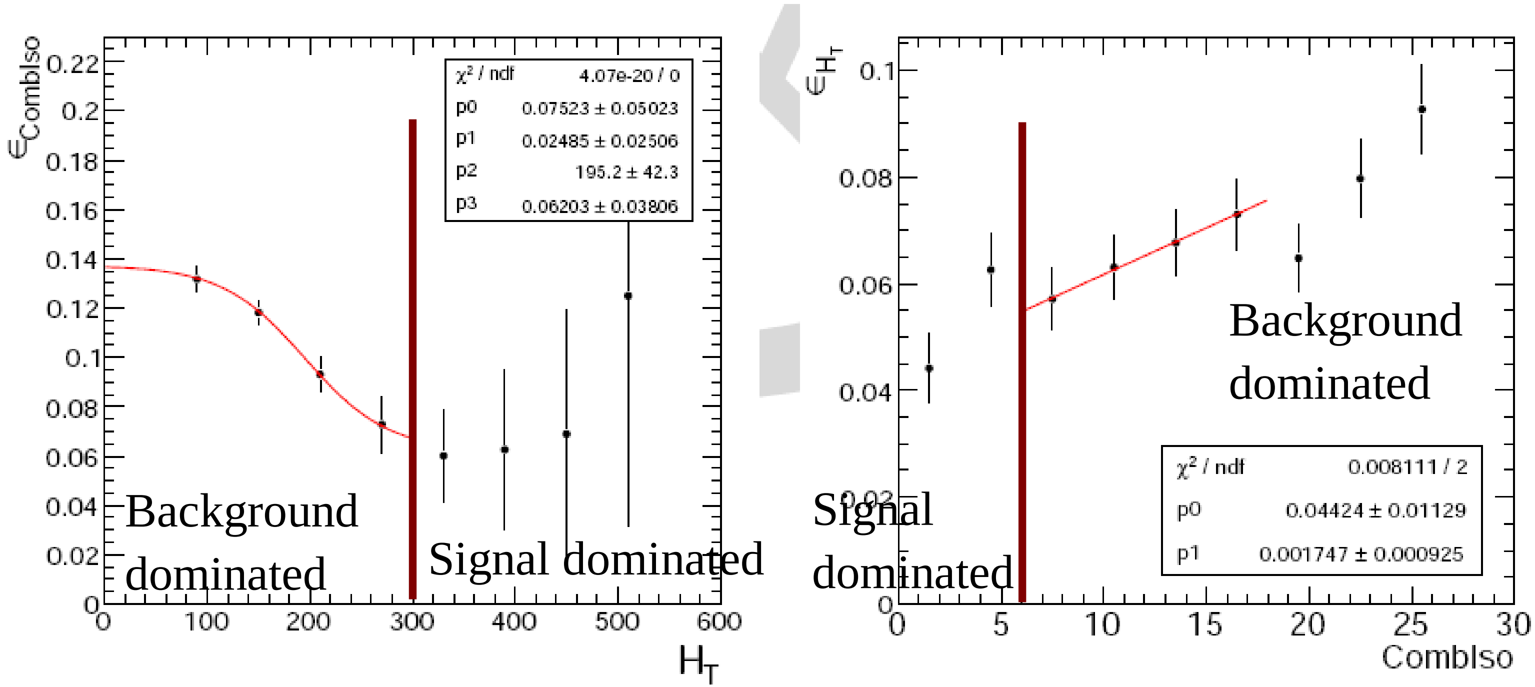
- Upgraded ABCD method :
  - Parametrize ε(A->B) as a function of Y (take the correlation into account from data!).
  - Parametrization of the efficiency :
    - ε(A->B) is obtained from measured ε(C->D) in various Y windows in the background dominated region
    - Fit a curve ε(C->D) on the data points
    - Extrapolate the curve in the « signal region » (here ; A&B regions)
    - Calculate ε(A->B) with the extrapolated ε :



$$\frac{N_A}{N_B} = \frac{1}{y_0}\int_0^{y_0} \epsilon_{C\to D}(y)\,dy$$

Background Estimation from data

- Question : how well do we recover from correlations by fitting the curve?
- Example : CombIso vs H_T (to cross-check the results of CombIso vs p_T^{μ} for example)
  - Fit made in the region dominated by background.



Jet selection

- Question : « What are the cuts on the jet E_T that minimize the tt+jets cross section uncertainty, assuming X% of systematic uncertainty on the number of selected QCD events » ?

- This study does aim at :
  - Showing that asymmetrical cuts should be envisaged
- This study does not aim at :
  - Finding optimized cuts on the jet E_T as only the QCD background has been taken into account.

Technicalities

- CMSSW 1_6_12, PAT
- Background and signal sample :
  - QCD PYTHIA "/ppMuPt20-15/CMSSW_1_6_7-CSA07-1204357236/RECO" :
    - ~2M events = 8.7pb^{-1} (rescaled to 10pb^{-1})
    - 100 pb^{-1} misalignment, miscalibration scenario
    - HLT1MuonNonIso NO trigger applied.
  - Chowder alpgen : tt+jets (and W/Z+jets)
    - Skimmed according to HLT1MuonNonIso + 1stJet E_T>65GeV, rescaled to 10pb^{-1}
- Event pre-selection :
  - At least one reconstructed muon with p_T>20 GeV/c, |eta|<2.4
  - At least four reconstructed jets with E_T>30 GeV, |eta|<2.4
  - RelIso >0.95 and ΔR(μ,closest jet among the 4 leading ones)>0.3

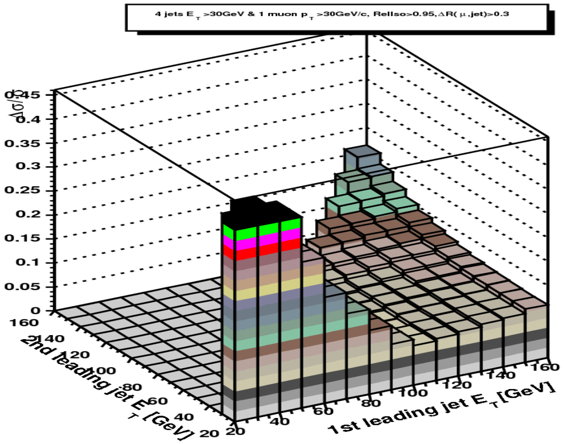
tt+jets cross section uncertainty

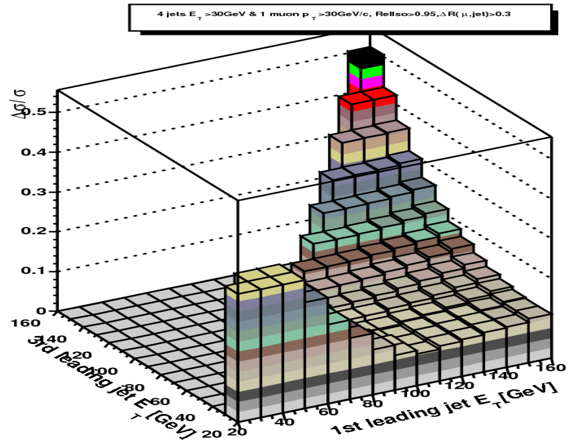
- Assuming :
  - QCD background only
  - No error on the integrated luminosity
  - No error on the selection efficiency (not relevant if derived from MC)
  - Poisson uncertainty on the number of selected events
  - X% of systematic uncertainty on the number of selected QCD events

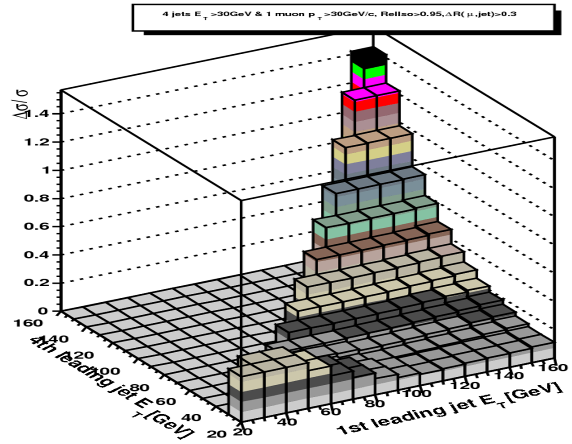
$$\sigma_{t\bar{t}} = \frac{N_{tot}^{selected} - N_{QCD}^{selected}}{\epsilon^{selection} \int \mathcal{L}}$$

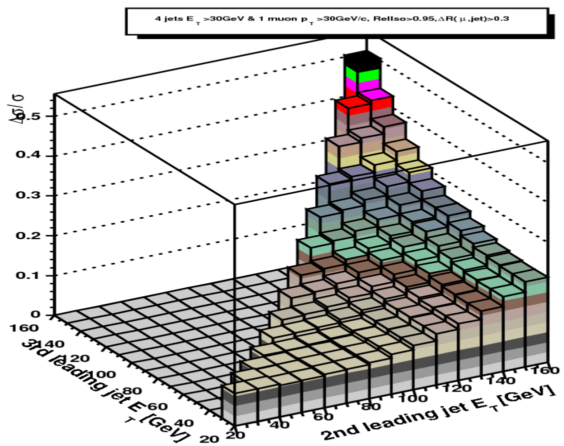
$$\frac{\Delta\sigma_{t\bar{t}}}{\sigma_{t\bar{t}}} = \frac{\Delta\left(N_{tot}^{selected} - N_{QCD}^{selected}\right)}{N_{t\bar{t}}^{selected}} = \frac{\sqrt{\Delta(N_{tot}^{selected})^2 + \Delta(N_{QCD}^{selected})^2}}{N_{t\bar{t}}^{selected}} = \frac{\sqrt{N_{tot}^{selected} + ((1 + (X/100)^2)N_{QCD}^{selected})^2}}{N_{t\bar{t}}^{selected}}$$

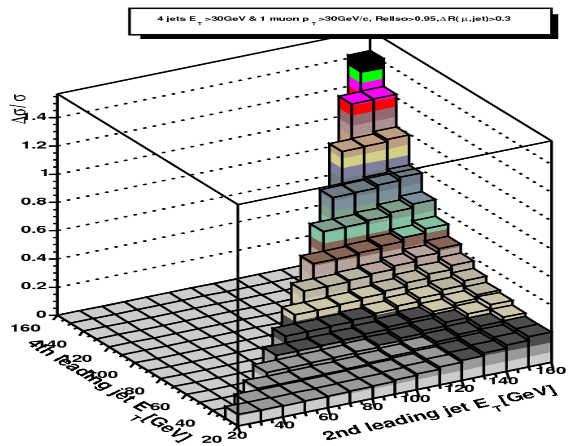
Jet selection cuts

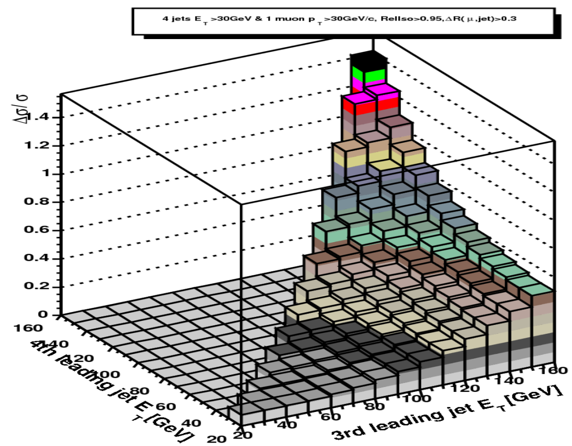












Jet selection cuts

	Pre-selection	QCD 118	tt+jets 313	W+jets 151	Z+jets 12
QCD systematics	Jet ET				
5%	90/50/40/30	7	206	76	6
		Cross section uncertainty = 7.8%			
30%	100/40/30/30	6	195	87	7
		Cross section uncertainty = 7.9%			
80%	90/40/40/40	1	145	41	3
		Cross section uncertainty = 8.5%			

- With RelIso, it possible to reject the « QCD background » (to be checked with newer samples...)
- But RelIso can not be used with the ABCD method...
- Sharing the same selection criteria across the Top group makes life easier but selection criteria depend on the analysis aim...