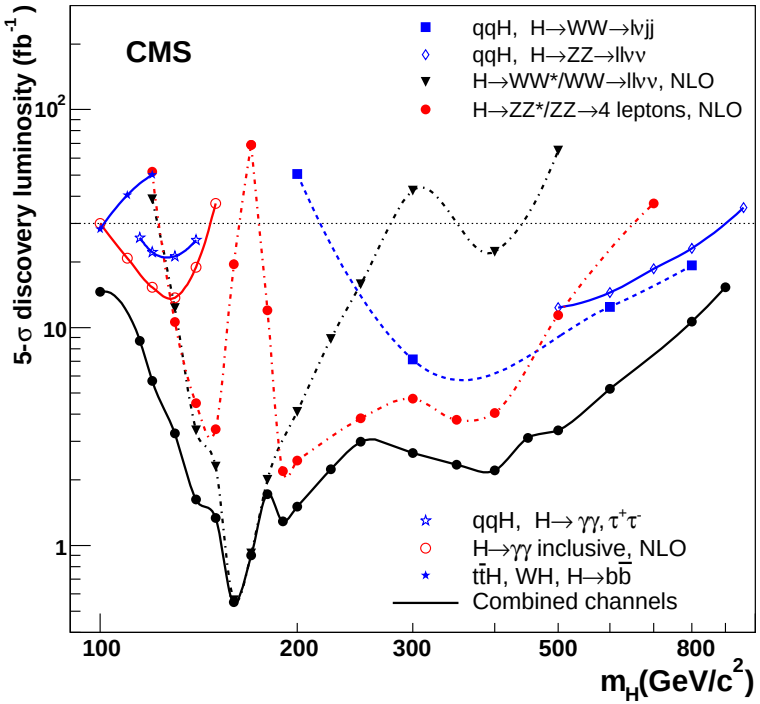
Multi-jet backgrounds to a heavy Higgs boson at the LHC

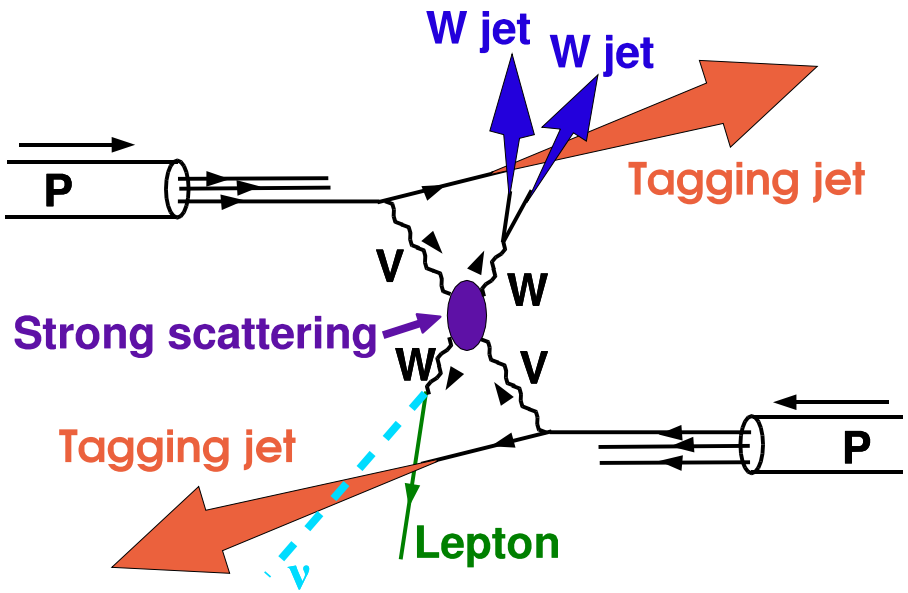
Paweł Zych

IEP Warsaw  CMS group

March 1, 2006

Searches for a heavy Higgs boson at the LHC





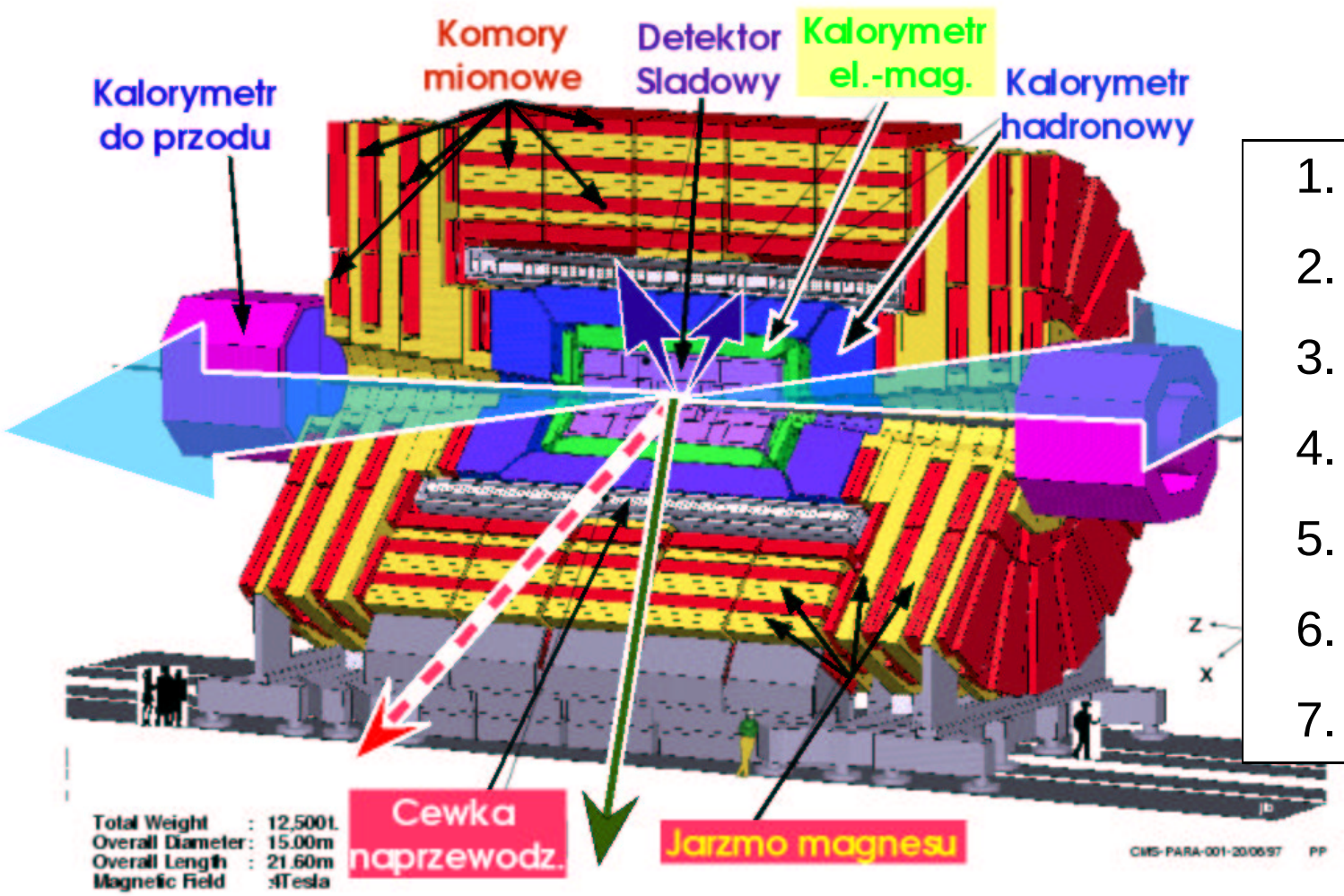
heavy Higgs boson/ strongly-coupled Higgs sector:

- production: VBF
- decay: H→WW→ $l\nu q\bar{q}$

My searches

1. W→ $\mu\nu$
   - hard, central and isolated muon
   - hard $\nu \Rightarrow$ large MET
2. W→$q\bar{q}$
   - 2 (1) very hard and central jets
3. 2 hard jets in forward regions (tagging)

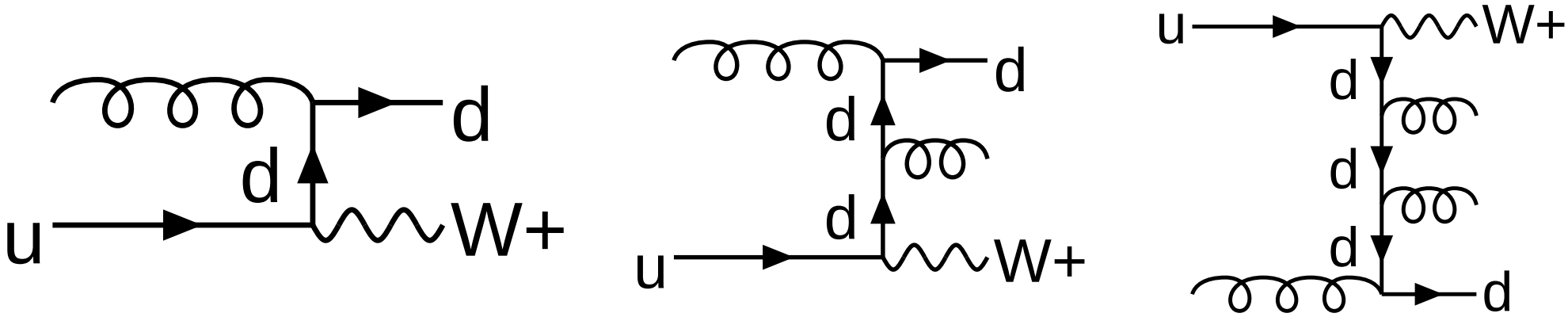
Selection procedure



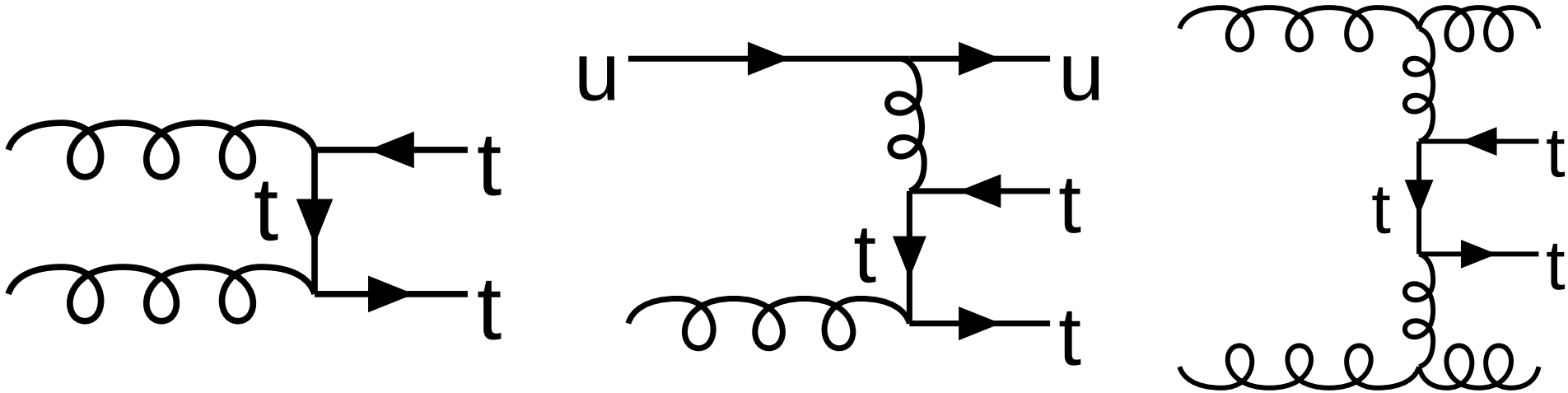
1. # of jets and μ
2. W$\rightarrow \mu\nu$ reconstruction
3. W$\rightarrow q\bar{q}$ reconstruction
4. top veto
5. tagging jets
6. pT(WWjj)<50 GeV
7. mini jet veto

Considered multi-jet backgrounds

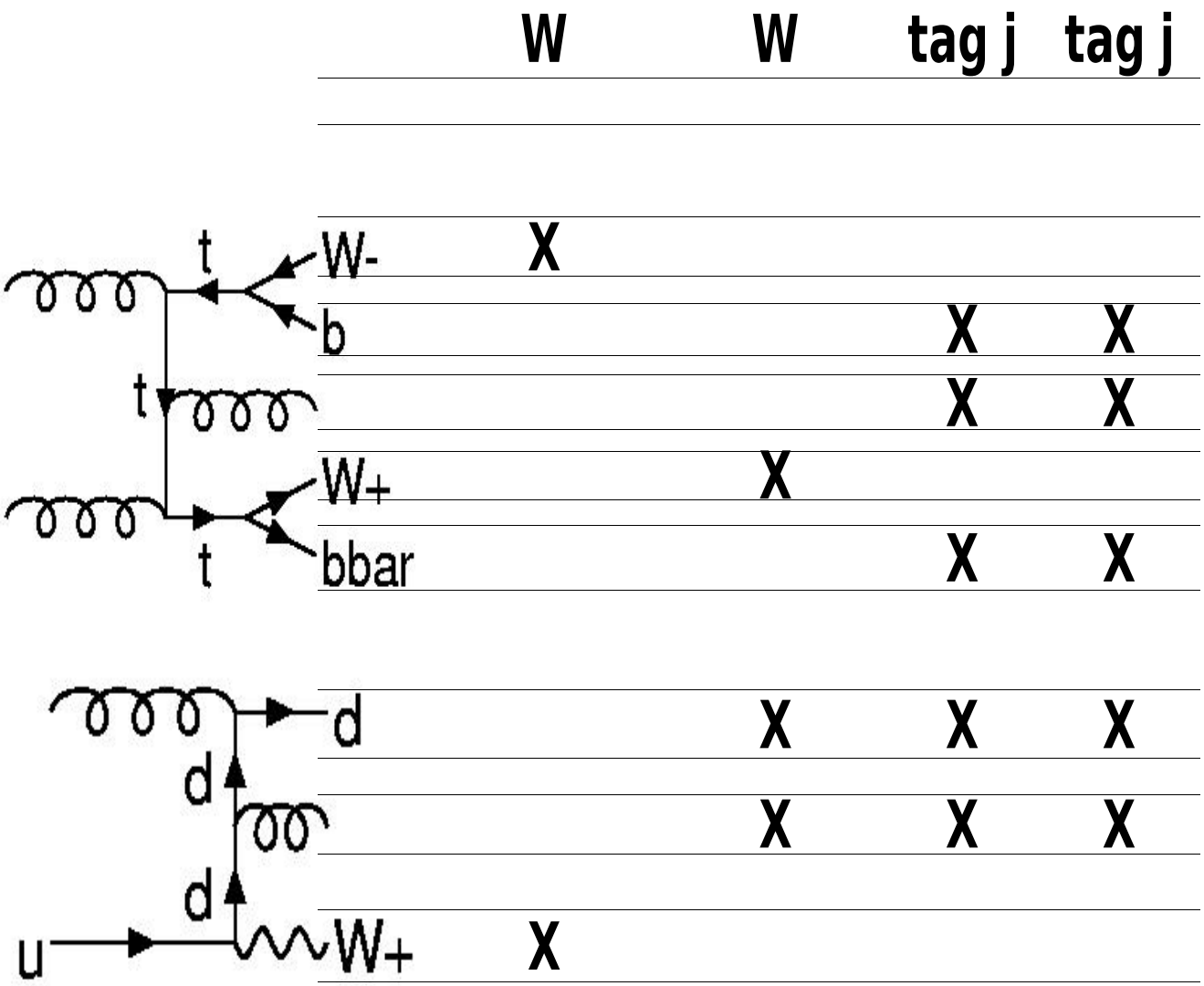
1. 1 real W + partons: Wj, Wjj, Wjjj (j=u, $\bar{u}$, d, $\bar{d}$, g)



2. $t\bar{t}$+ partons: $t\bar{t}$, $t\bar{t}j$, $t\bar{t}jj$ (j=u, $\bar{u}$, d, $\bar{d}$, g)

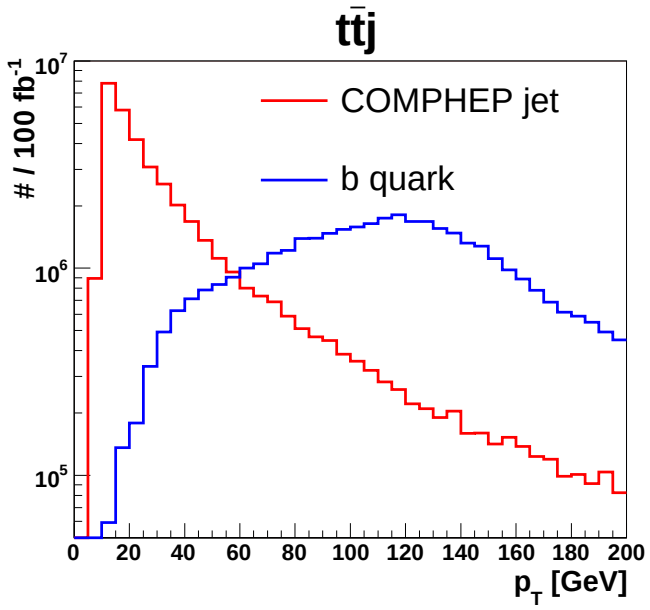


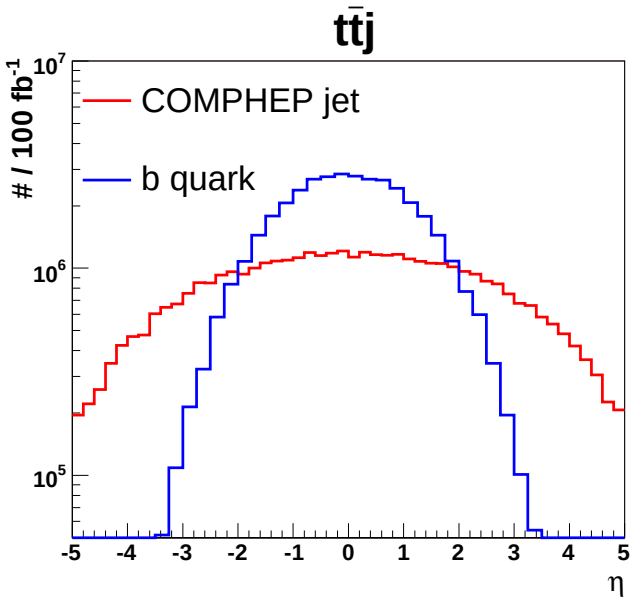
How $t\bar{t}$+jets and W+jets are the backgrounds



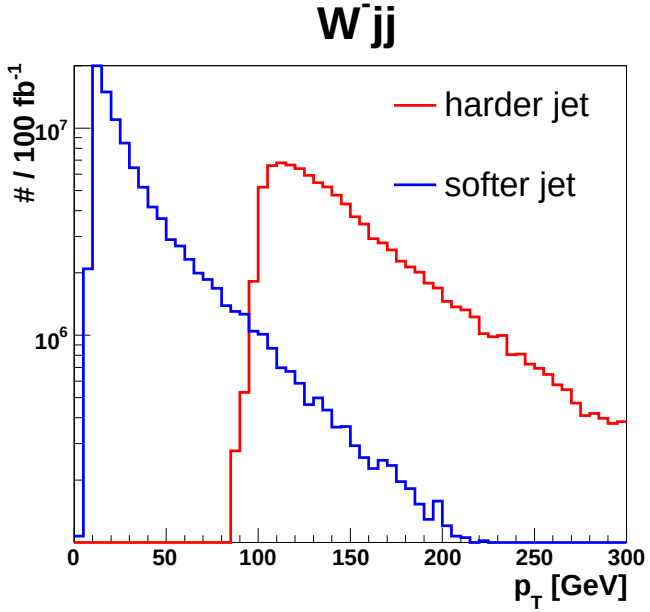
- W+jets: W$\rightarrow \mu\nu$
- $t\bar{t}$+jets:
  $t\bar{t}\rightarrow W(\rightarrow \mu\nu)W(\rightarrow q\bar{q})b\bar{b}$
- sources of tag jets:
  - jets at generation lev.
  - b quarks
  - ISR jets
  - influence of pile-up (PU)
  - pile-up
- more gener.lev. jets $\Rightarrow$ more important background

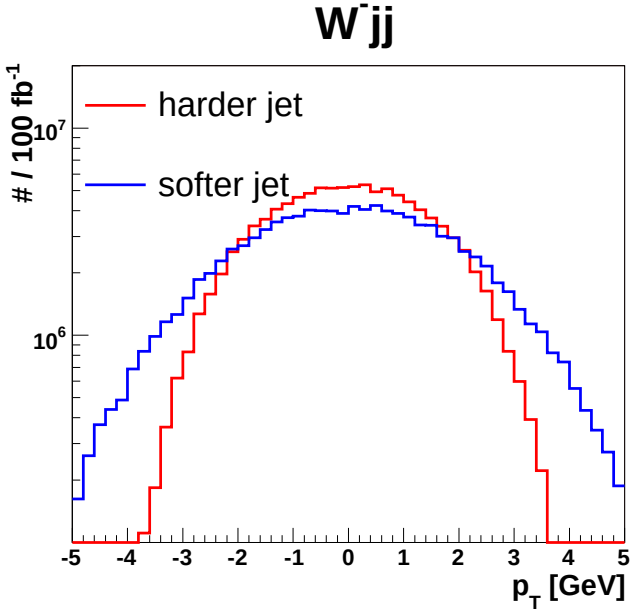
How $t\bar{t}$+jets and W+jets are the backgrounds-kinematics of jets





- gener.lev. jet is a better source of 'tag jet' than a b-jet
- b quark jets will be vetoed:
  b tag veto, top veto, mini jet veto





- harder jet is a natural source of hadronic W
- softer jet becomes a tag jet

Generation of $t\bar{t}$+jets and W+jets

- generation cuts:

$t\bar{t}$, $t\bar{t}$+j, $t\bar{t}$+jj

cut	min	max
pT for jets	10.	
pt fot t	100.	
eta for jets	-5.5	5.5
eta for t	-5.5	5.5
dR between j's	0.5	

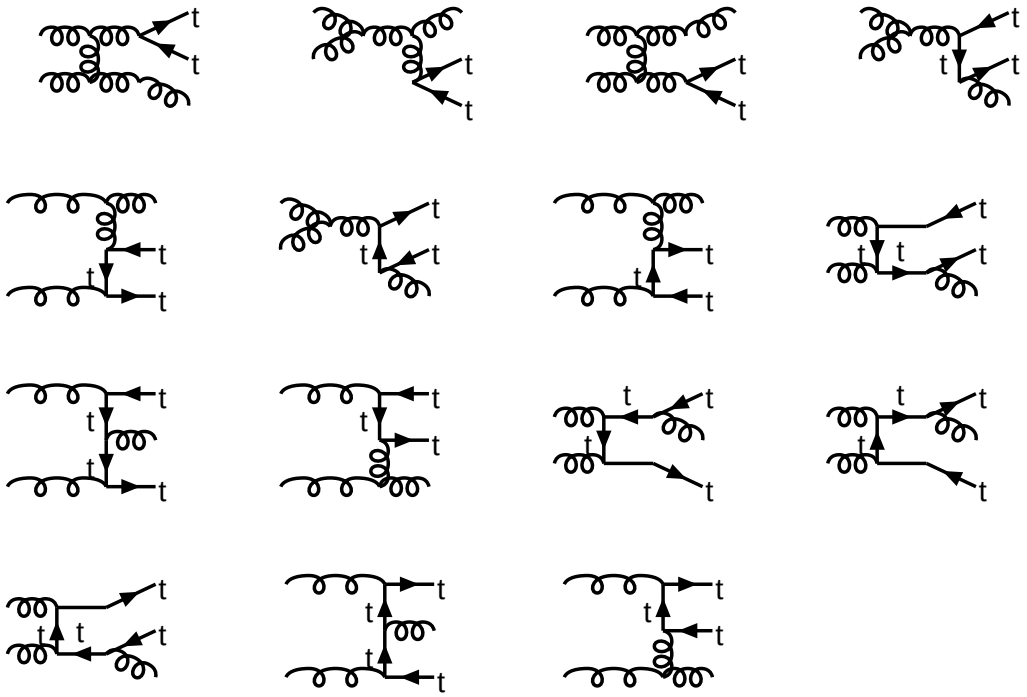
Wj, Wj+j, Wj+jj

cut	min	max
PT for j	10.	
highest PT for j	100.	
PT for W	100.	
eta for j	-5.	5.
eta for W	-3.	3.
dR between j's	0.2	

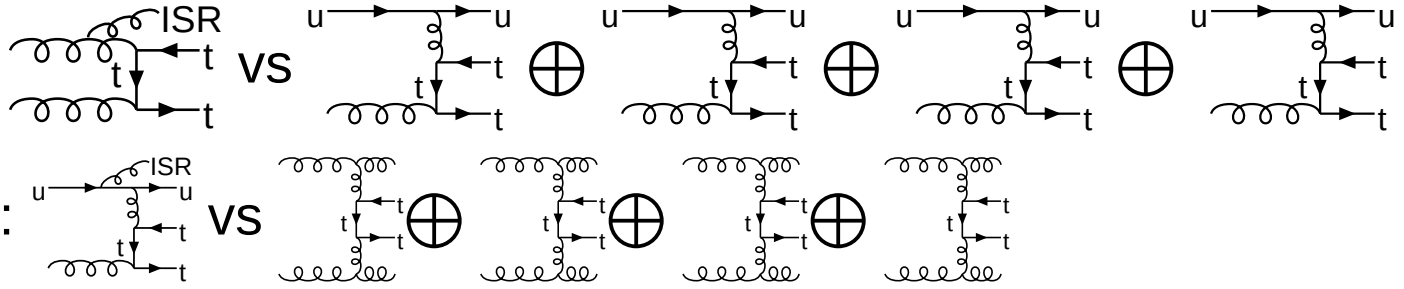
- Wj i $t\bar{t}$ can be generated in PYTHIA/COMPHEP
- Wjj, Wjjj, $t\bar{t}$j, $t\bar{t}$jj generated in steps:
  1. COMPHEP: generation (matrix elements)
  2. PYTHIA: ISR is added (j_{ISR}), fragmentation, hadronisation

The problem of possible overlap of the samples

- diagrams for $gg \to t\bar{t}g$:

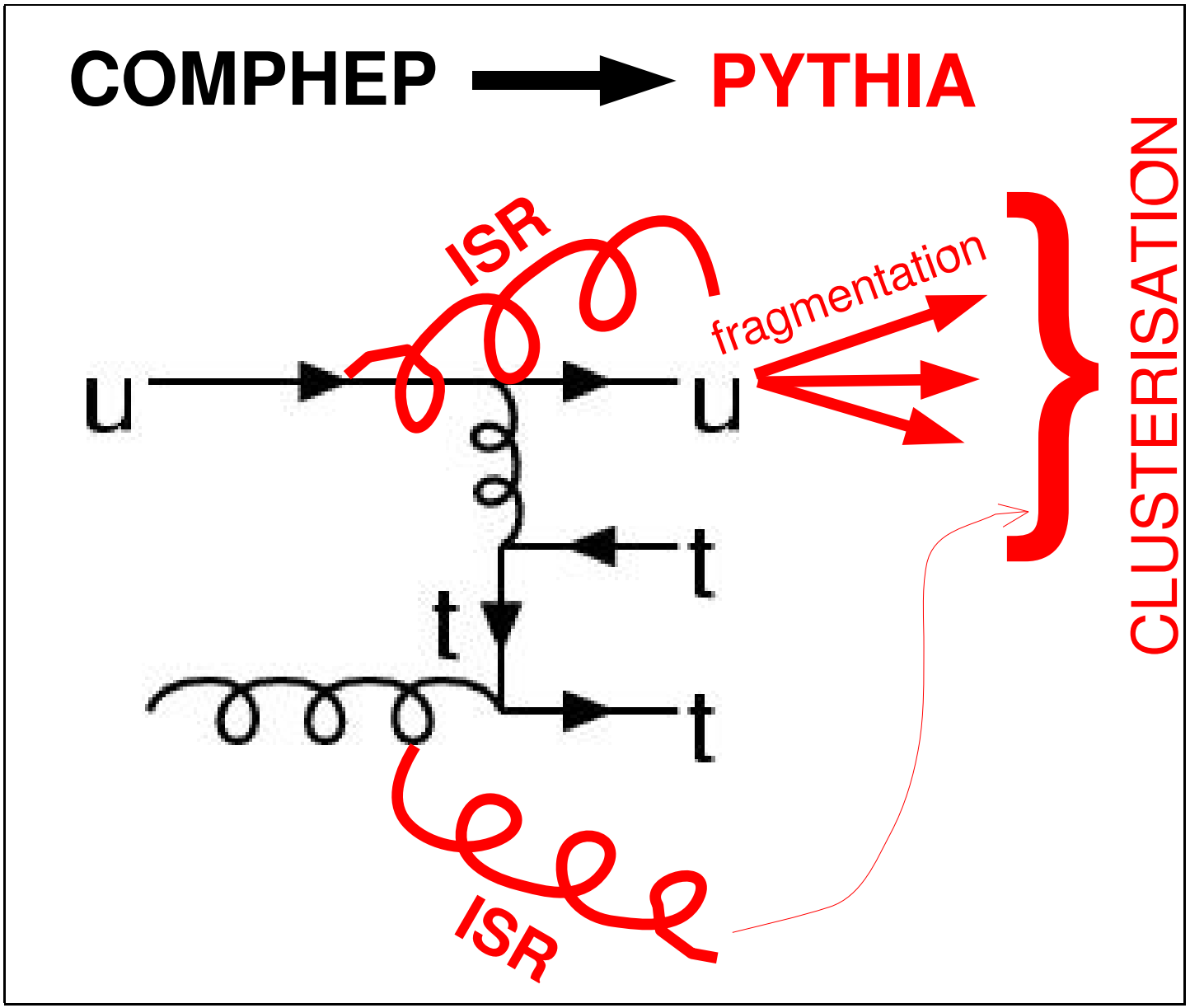


- ISR jets are both in COMPHEP and also added in PYTHIA:
  - $t\bar{t}+j_{ISR}$ vs $t\bar{t}j$:
  - $t\bar{t}j+j_{ISR}$ vs $t\bar{t}jj$:
  - ...



- double generation/overlap?

The method for the study of the problem



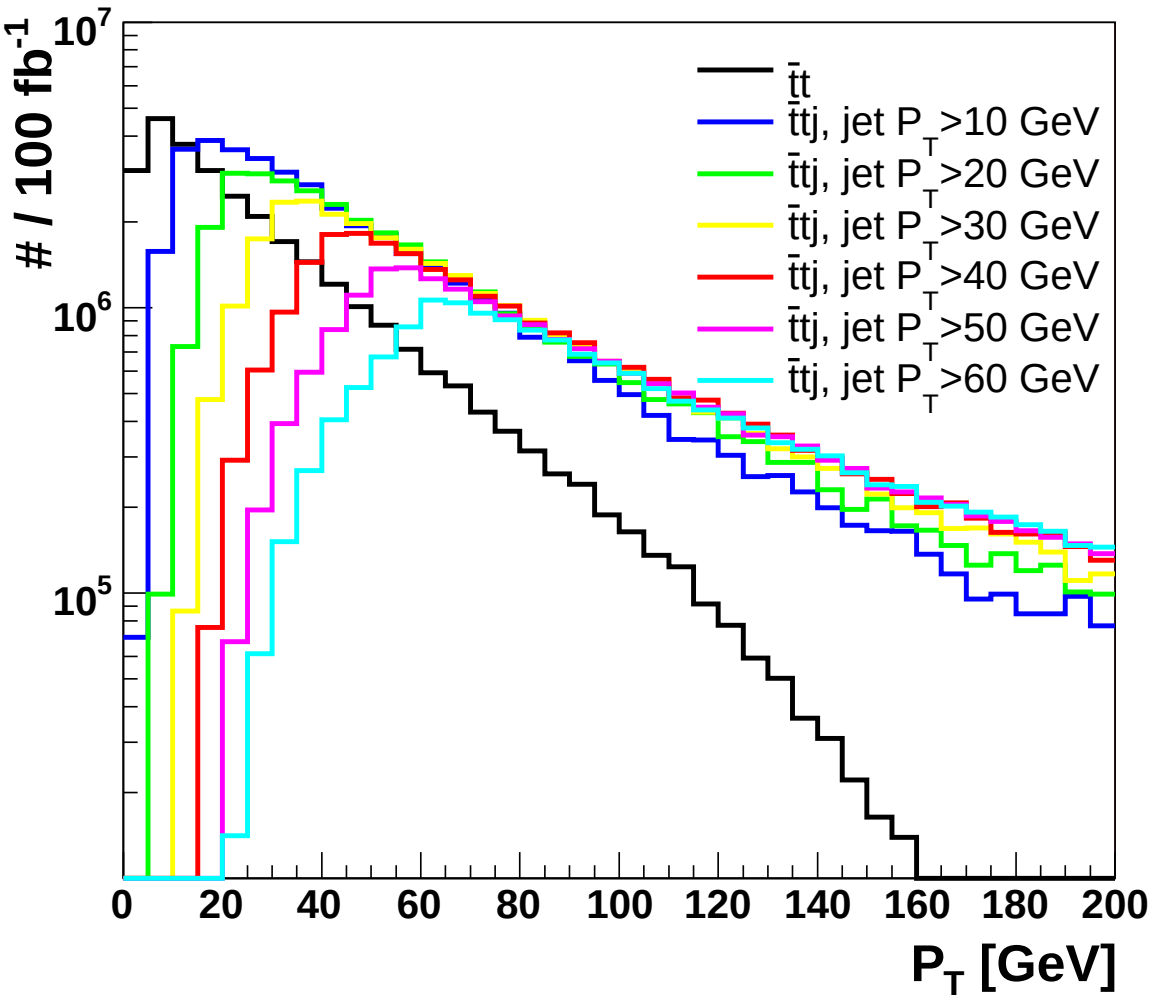
clusterisation:

- estimation of reconstructed jets at generation level
- partons are merged to groups (clusters)
- clustered:
  - ISR partons
  - fragmented j
- clusters: $p_{\perp}$>10 GeV

First approach to study the overlap of multi-jet samples

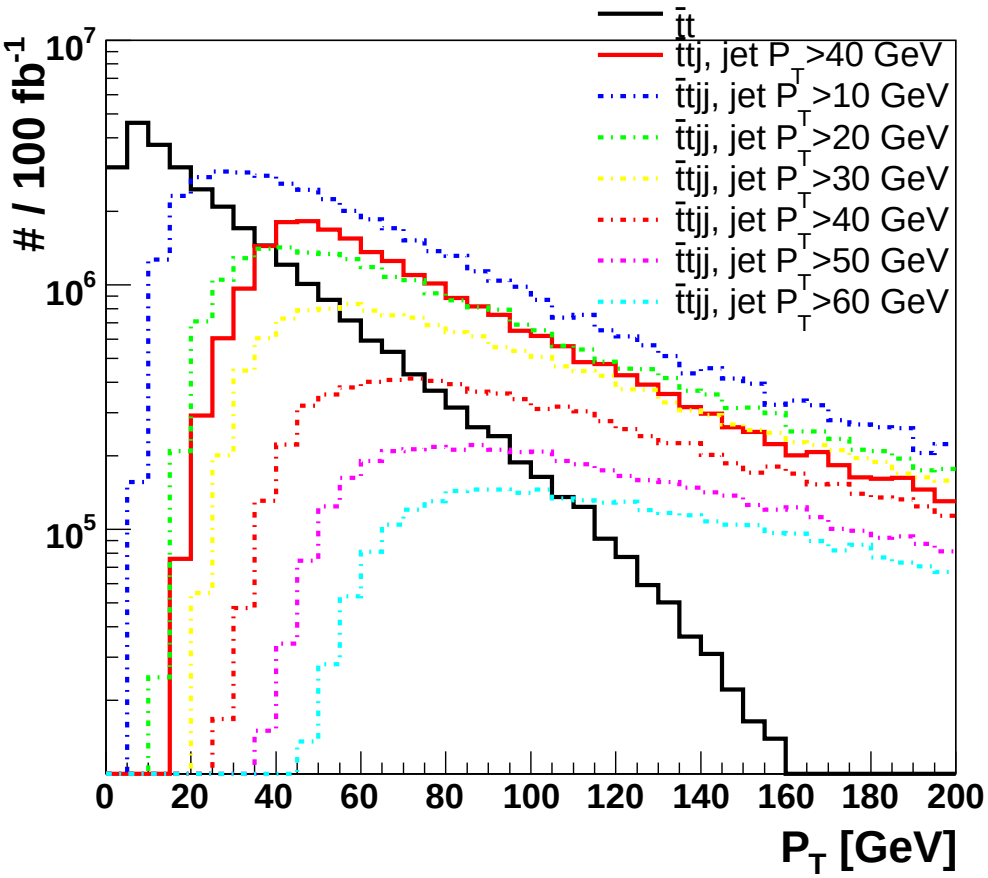
compare pT-distribution for the hardest cluster and so one for softer ones

hardest cluster for $t\bar{t}$ and $t\bar{t}$+j

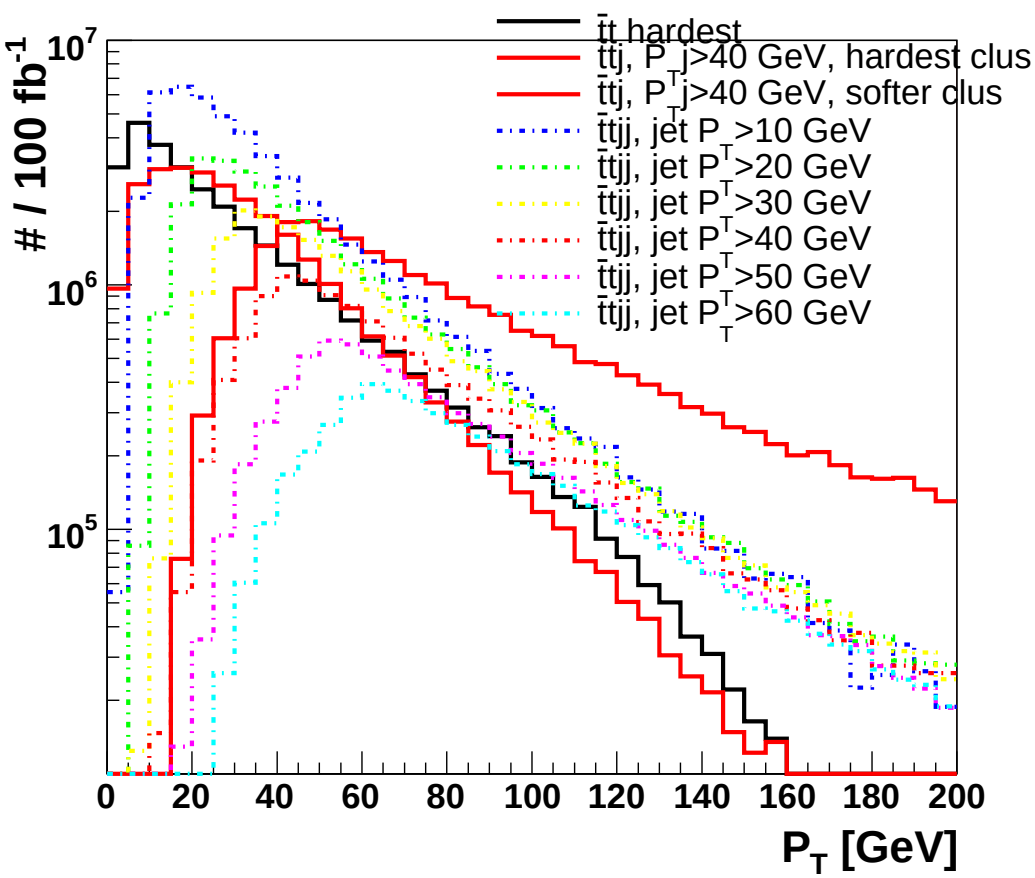


Clusters for $t\bar{t}j$ and $t\bar{t}jj$

hardest cluster



next to hardest cluster



Summary?

- we can only limit the overlap between the samples with different number of jets
- use the $p_\perp$ cuts on generated jets
  - $t\bar{t}j$: pT>40 GeV
  - $t\bar{t}jj$: pT$\gtrsim$100 GeV
  - Wjj: ...
  - Wjjj: ...
- profits:
  1. limited overestimation of the background
  2. σ's drops down with # of jets

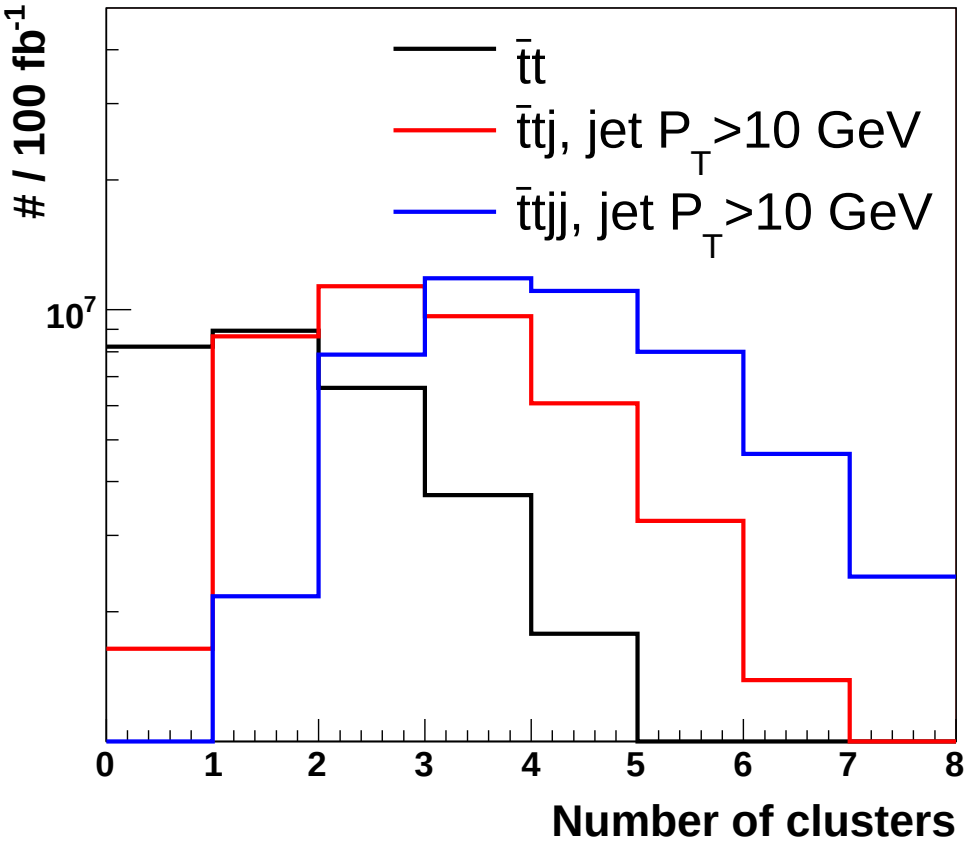
Cross sections [pb]

j pT cut	10 GeV	40 GeV	60 GeV
$t\bar{t}$	?? 600.		
$t\bar{t}$+j	428.	272.	190.
$t\bar{t}$+jj	494	116.	51.
Wmj	715.		
Wmj+j	1072.	363.	230.
Wmj+jj	973.	190.	94.
Wpj	1100.		
Wpj+j	1603.	561.	361.
Wpj+jj	1698.	346.	175.

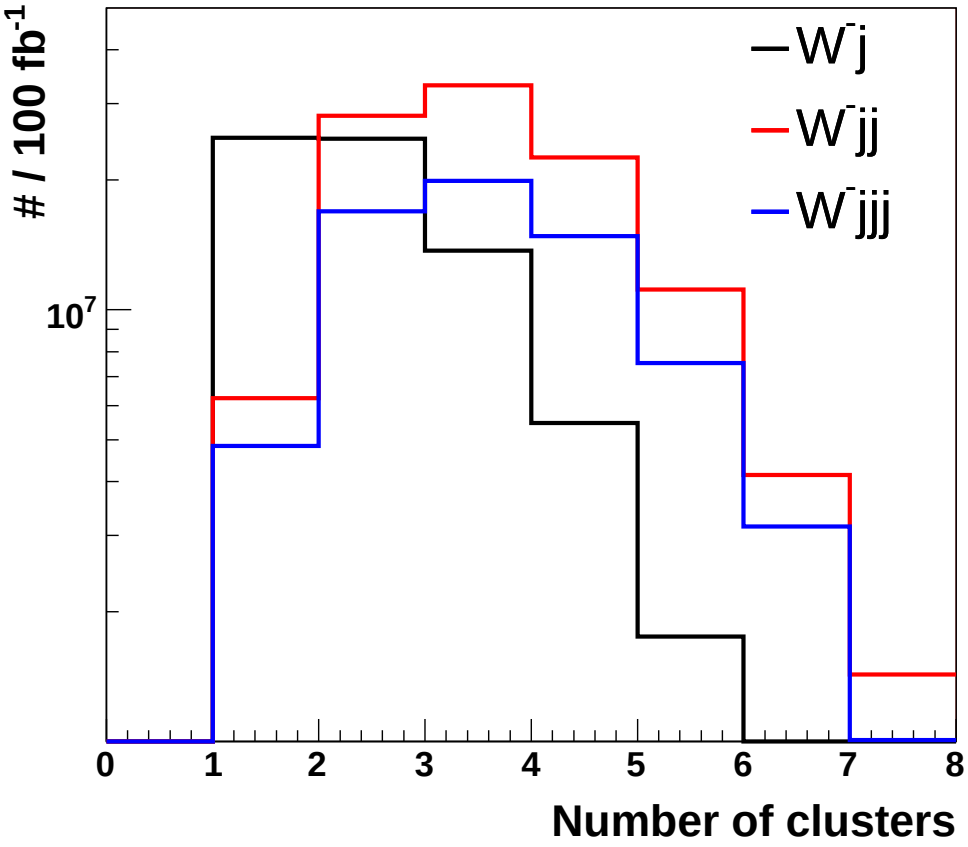
The other approach to study the overlap of multi-jet samples

comparison of pT-spectra for $t\bar{t}$+jets and W+jets samples with exactly the same number of clusters

clusters of pT>10 GeV

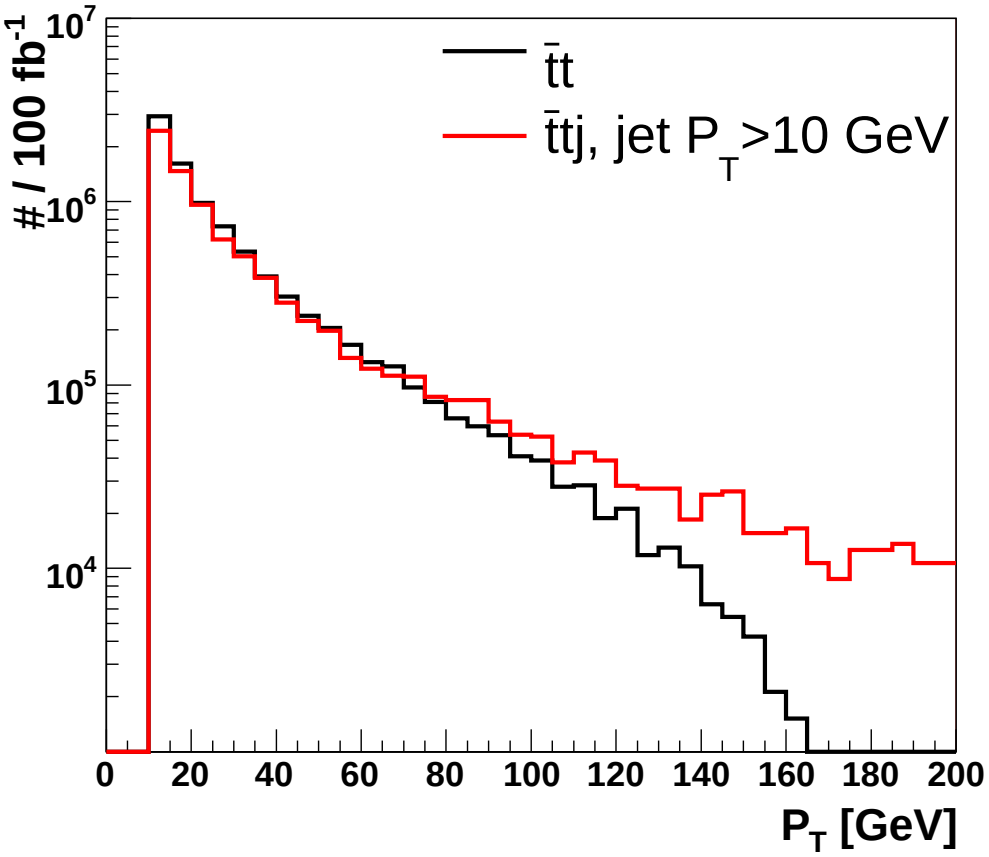


clusters of pT>10 GeV

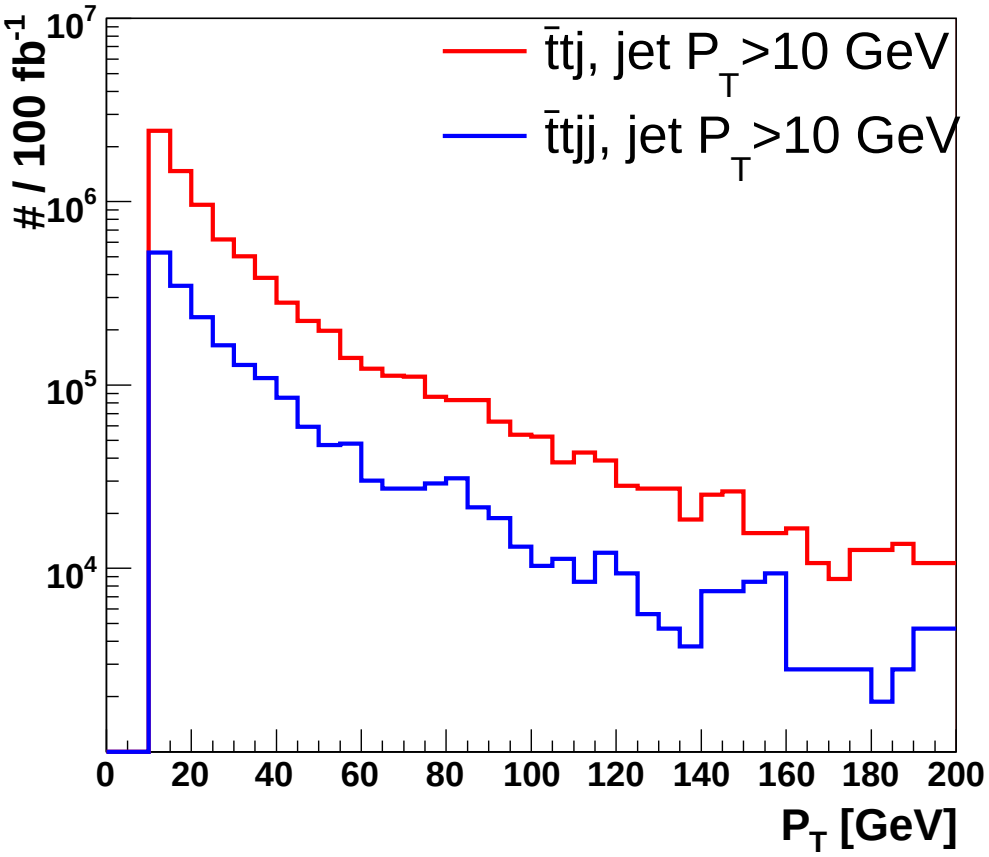


tt+jets, # clusters=1

$t\bar{t}$ and $t\bar{t}$+j

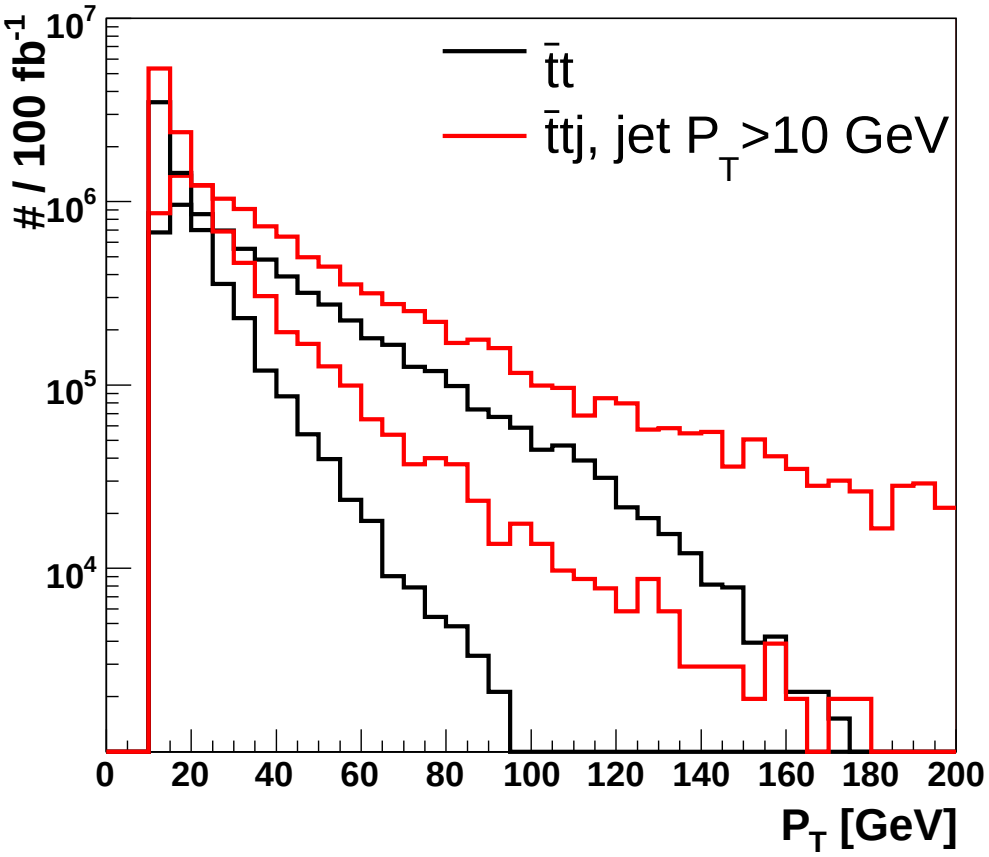


$t\bar{t}$+j and $t\bar{t}$+jj

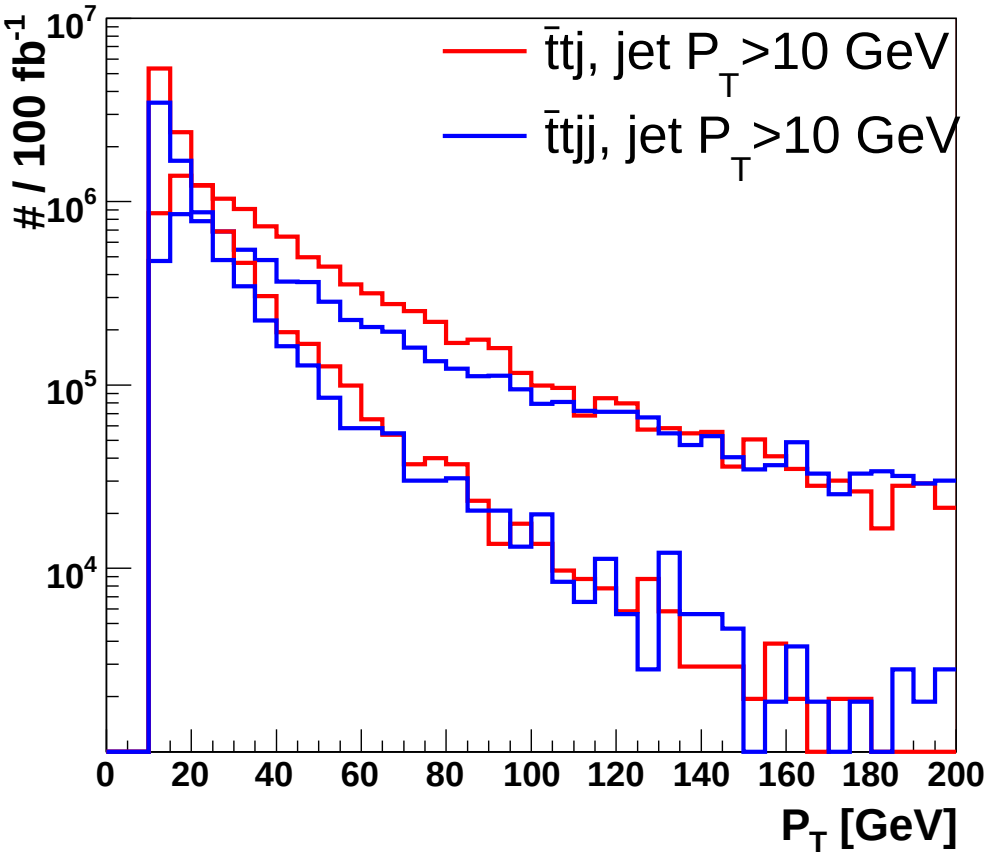


tt+jets, # clusters=2

$t\bar{t}$ and $t\bar{t}$+j

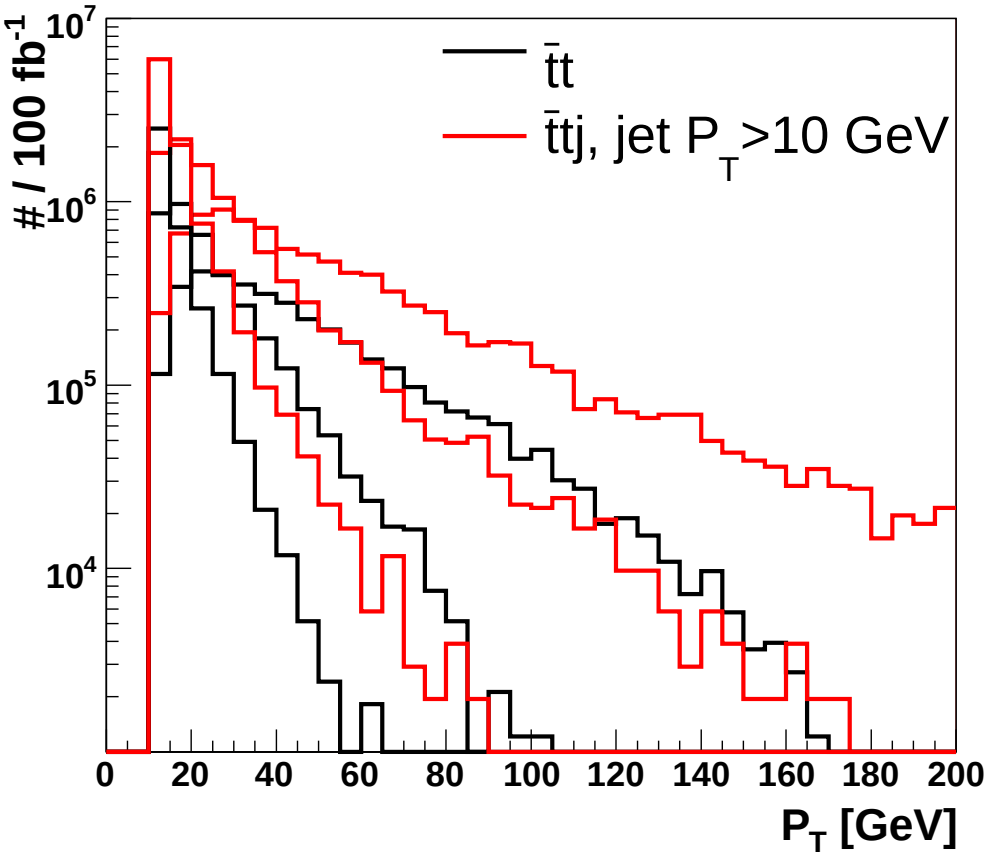


$t\bar{t}$+j and $t\bar{t}$+jj

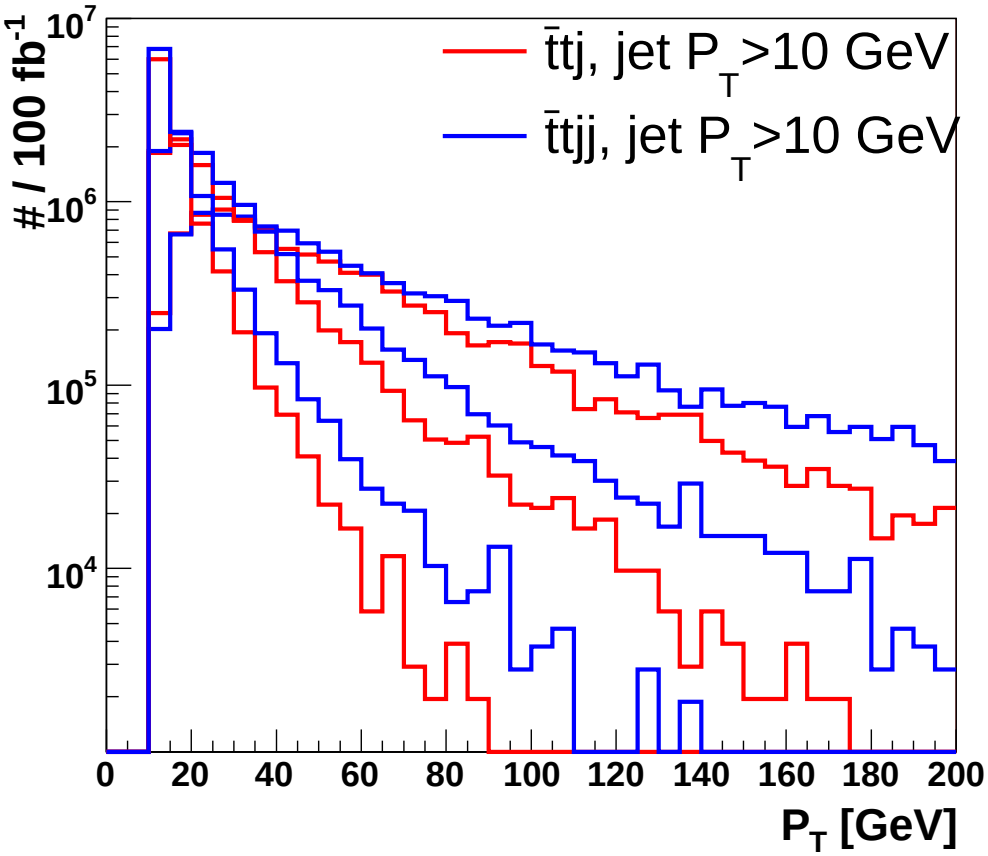


tt+jets, # clusters=3

$t\bar{t}$ and $t\bar{t}$+j



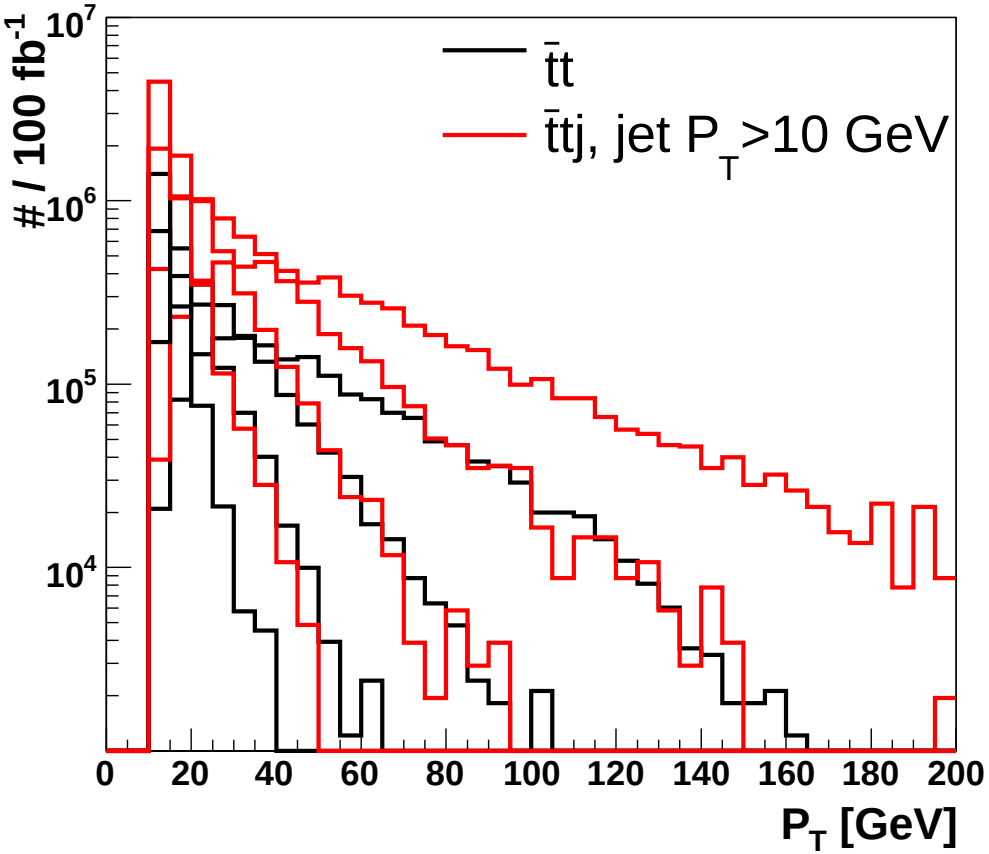
$t\bar{t}$+j and $t\bar{t}$+jj



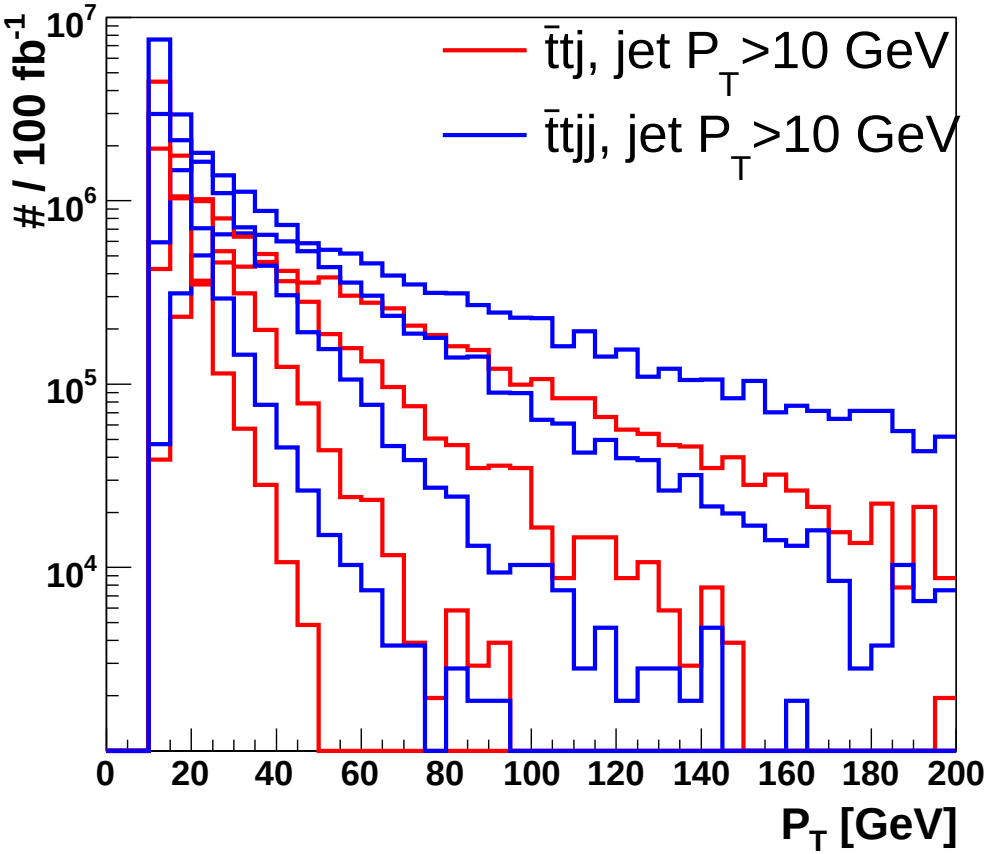
we do not need so many hard clusters (veto in selection!)

tt+jets, # clusters=4

$t\bar{t}$ and $t\bar{t}$+j



$t\bar{t}$+j and $t\bar{t}$+jj



we do not need so many hard clusters (veto in selection!)

Summary for $t\bar{t}$+jets

1. W(jj)W($\mu\nu$)jj signal $\Rightarrow$ not worth to look at # clusters$\geq$3

overlap between $t\bar{t}$ and $t\bar{t}$+j

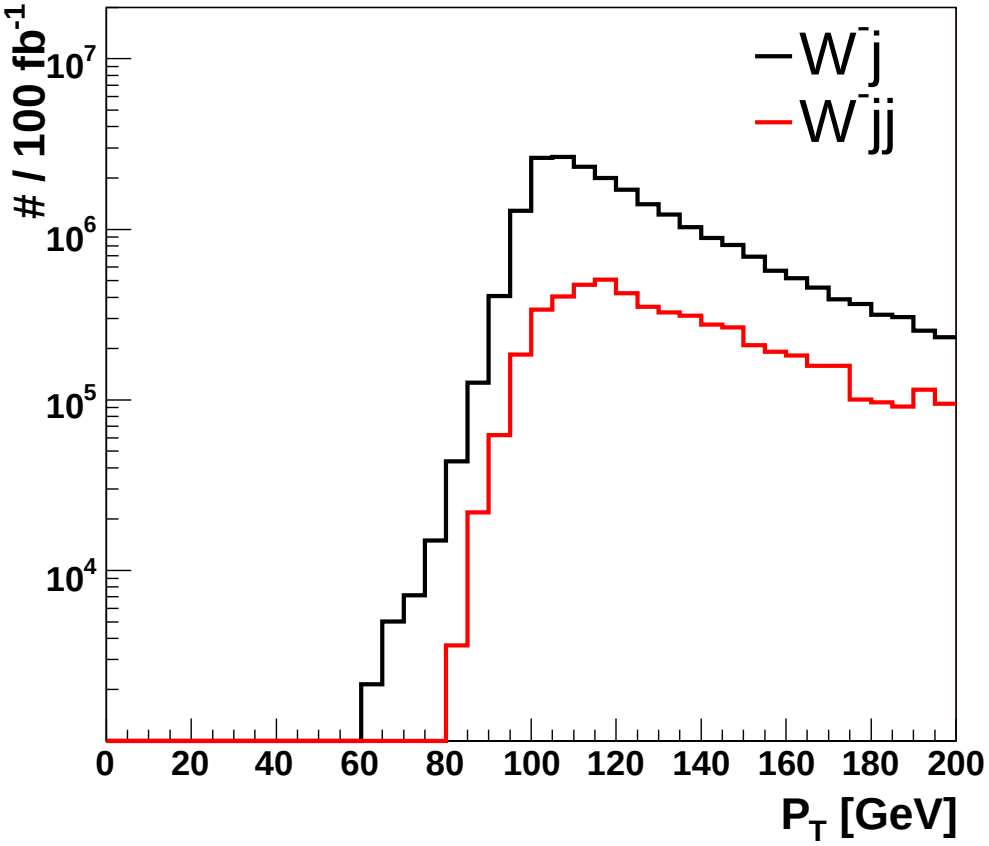
- # clusters=0 - not important
- # clusters=1 - diff. for $p_\perp > 140$ GeV
- # clusters=2 - difference
- # clusters$\geq$3 - important difference
- for this study: use $t\bar{t}$+j

overlap between $t\bar{t}$+j and $t\bar{t}$+jj

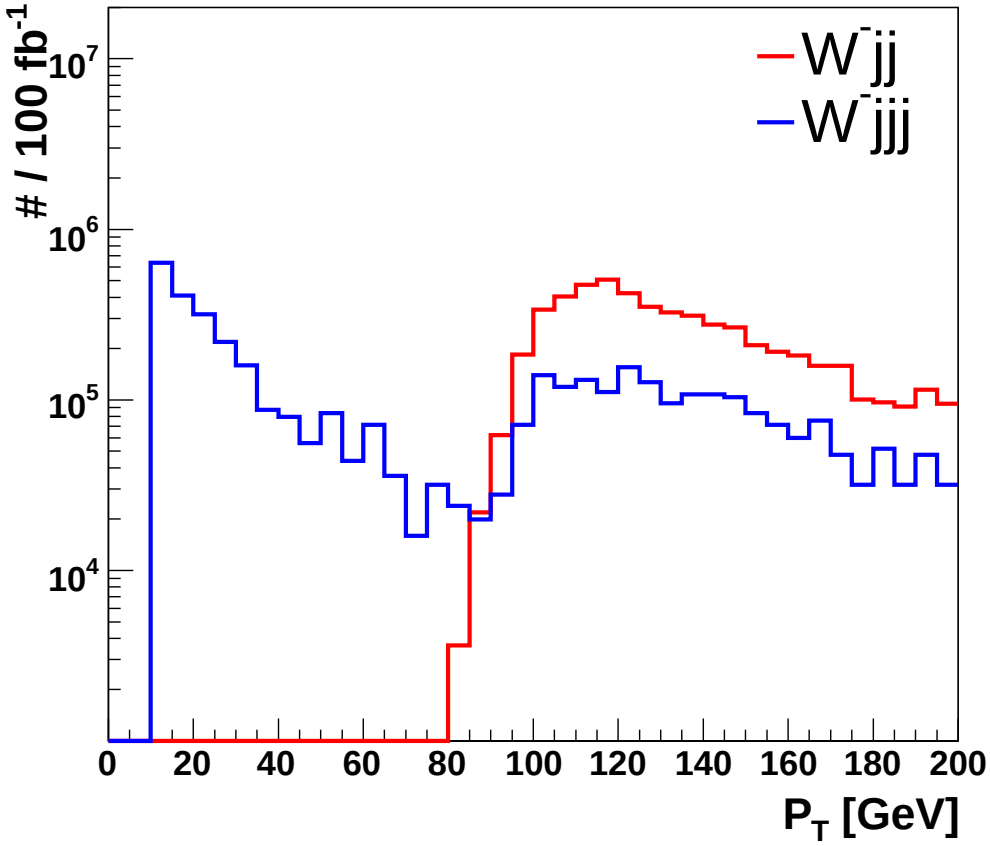
- # clusters=0 - not important
- # clusters=1 - problem with normalization, shape the same
- # clusters=2 - the same
- # clusters=3 - shapes the same, diff. in normalisation
- # clusters$\geq$4 - important difference
- for this study: use $t\bar{t}$+j

Wm+jets, # clusters=1

Wj and Wjj

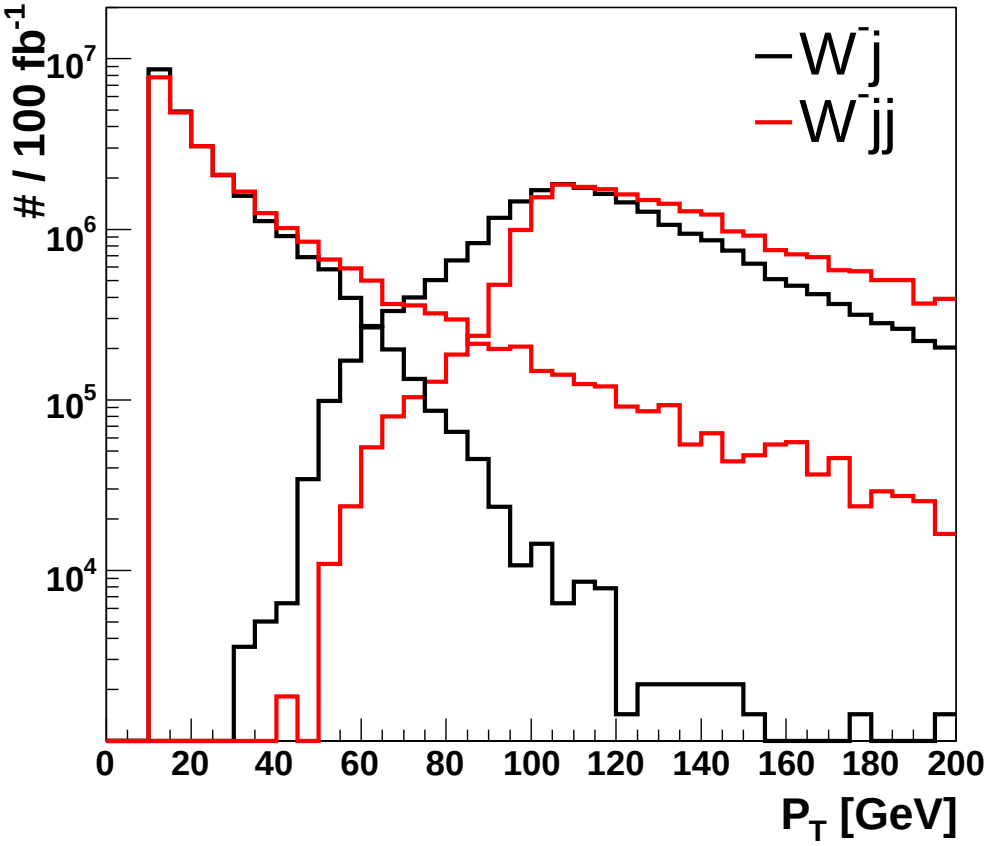


Wjj and Wjjj

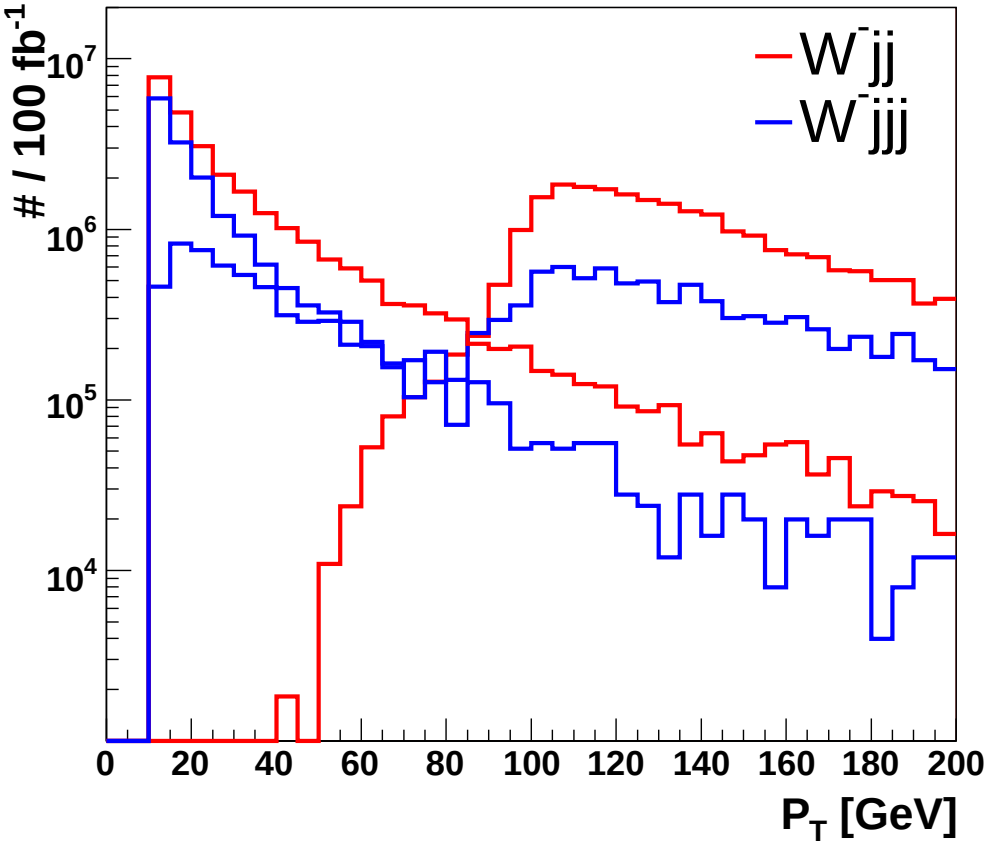


Wm+jets, # clusters=2

Wj and Wjj

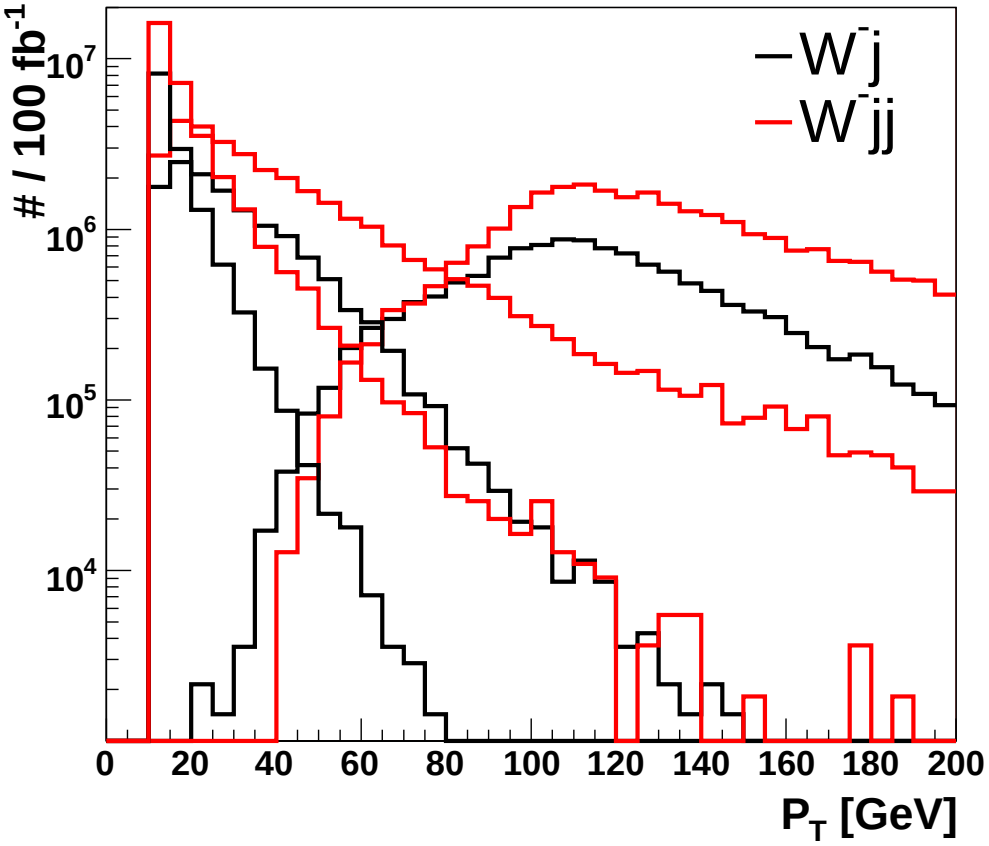


Wjj and Wjjj

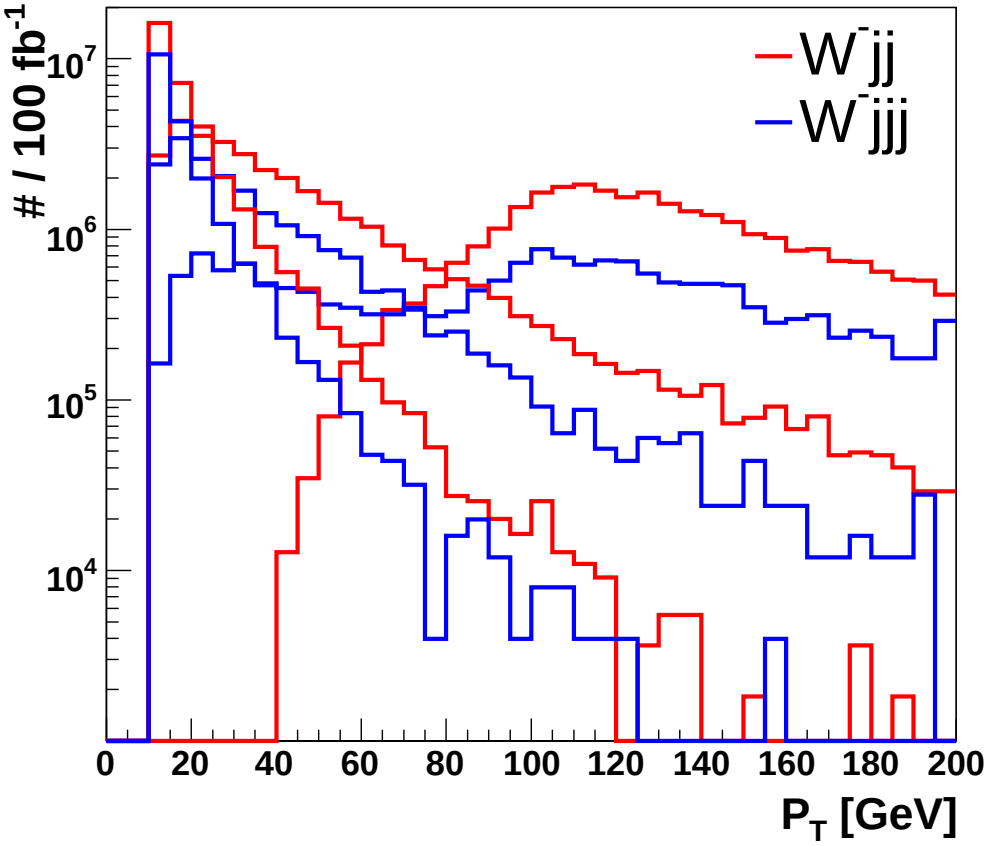


Wm+jets, # clusters=3

Wj and Wjj

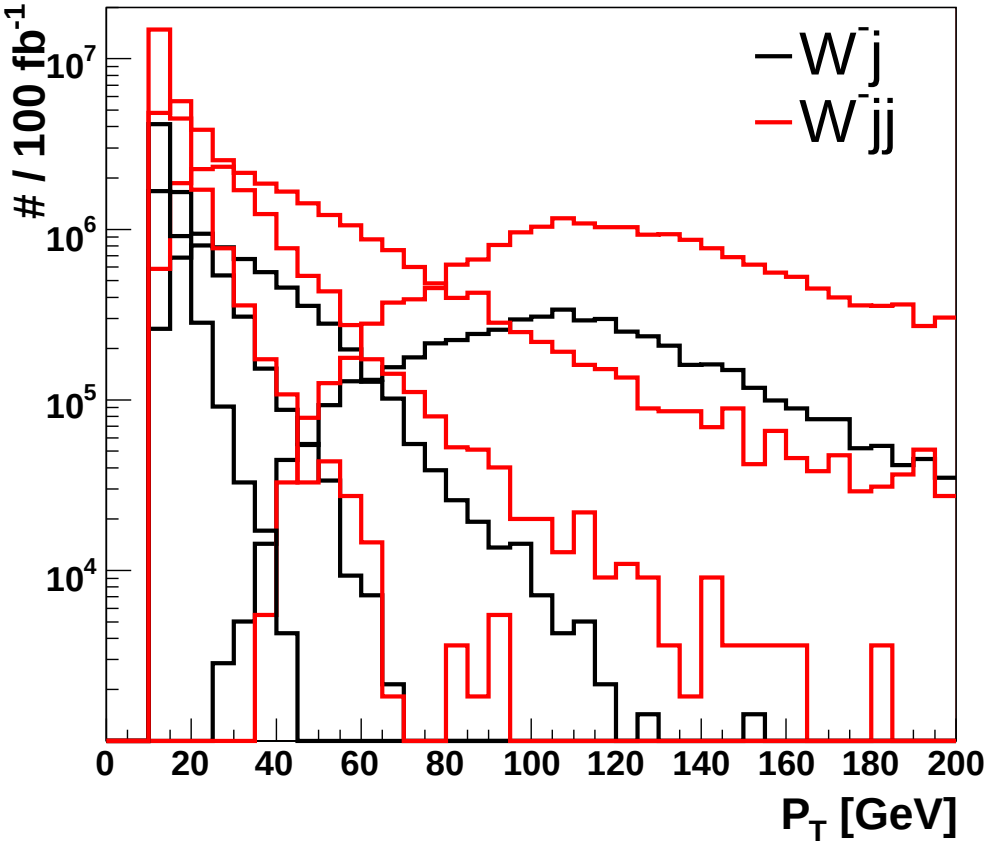


Wjj and Wjjj

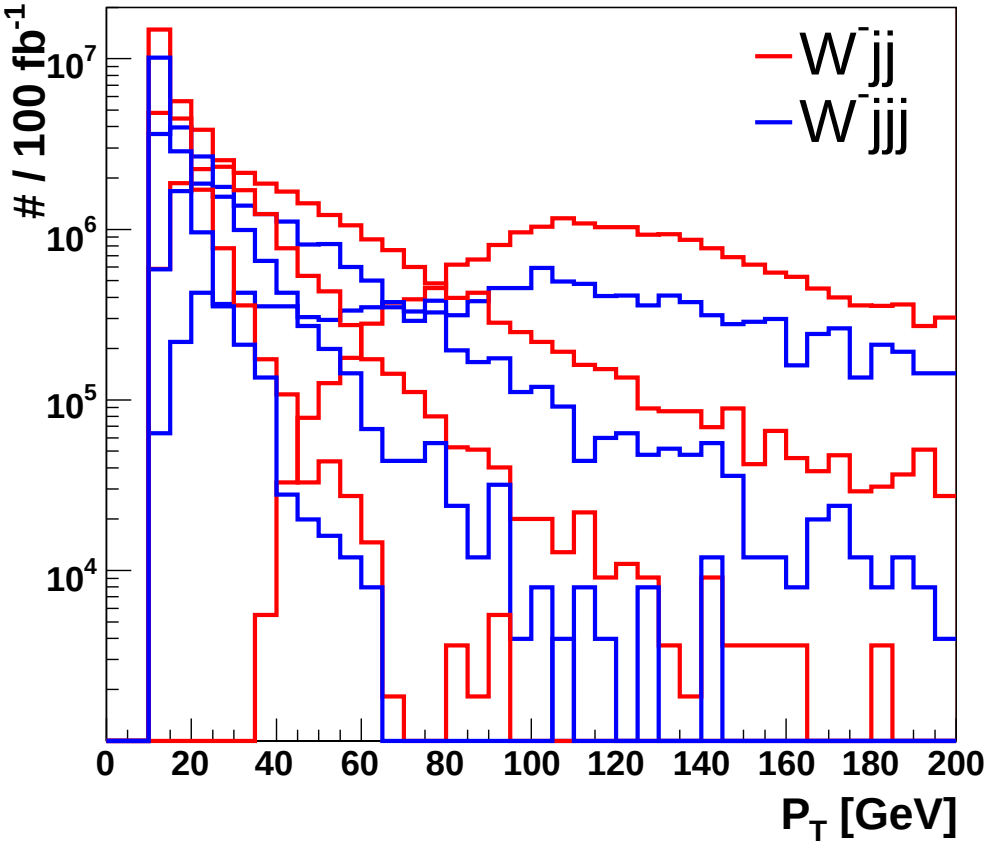


Wm+jets, # clusters=4

Wj and Wjj

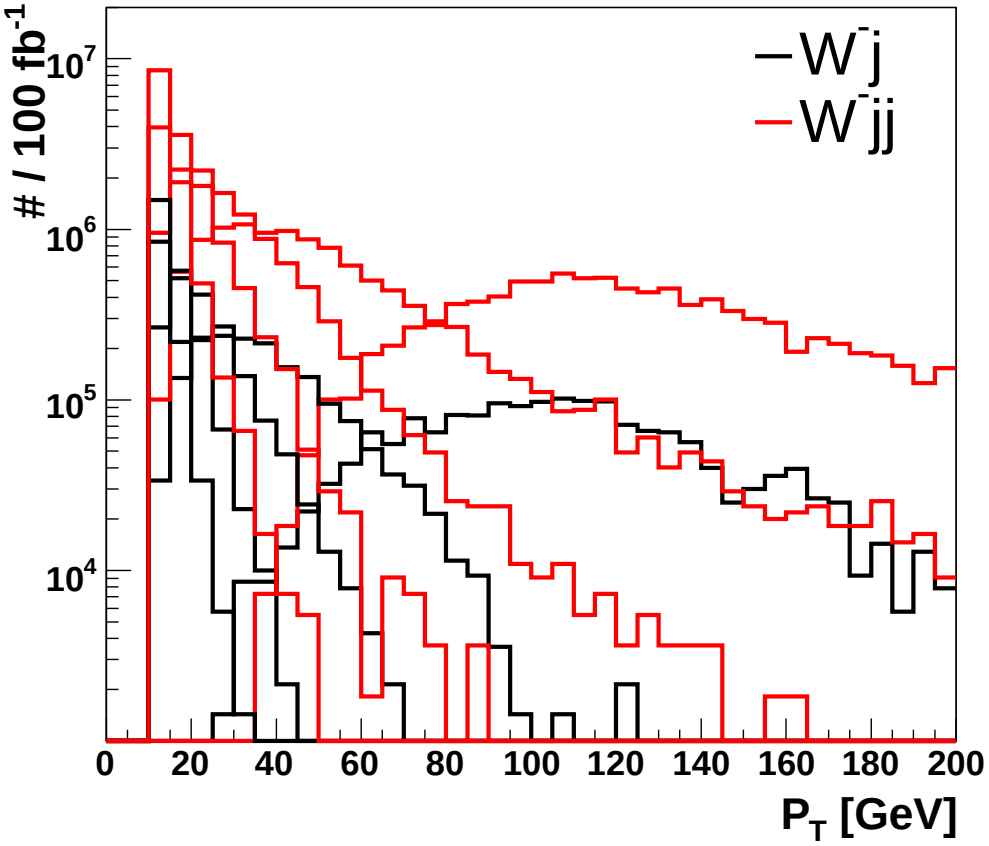


Wjj and Wjjj

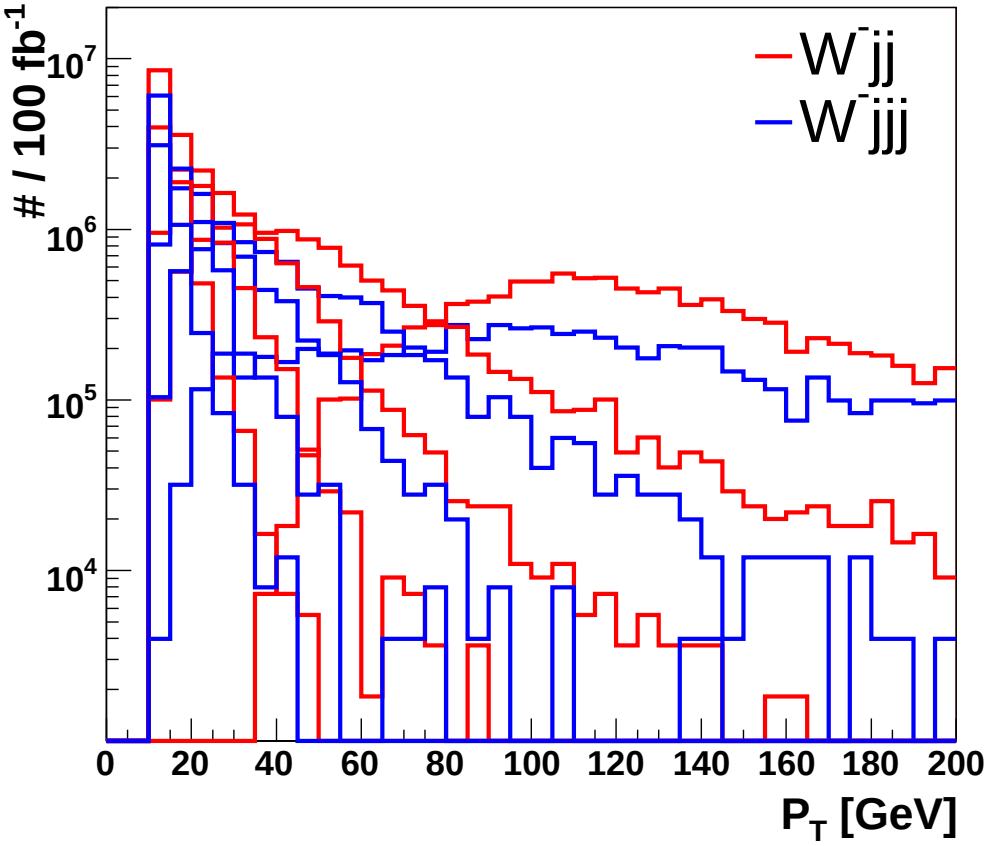


Wm+jets, # clusters=5

Wj and Wjj

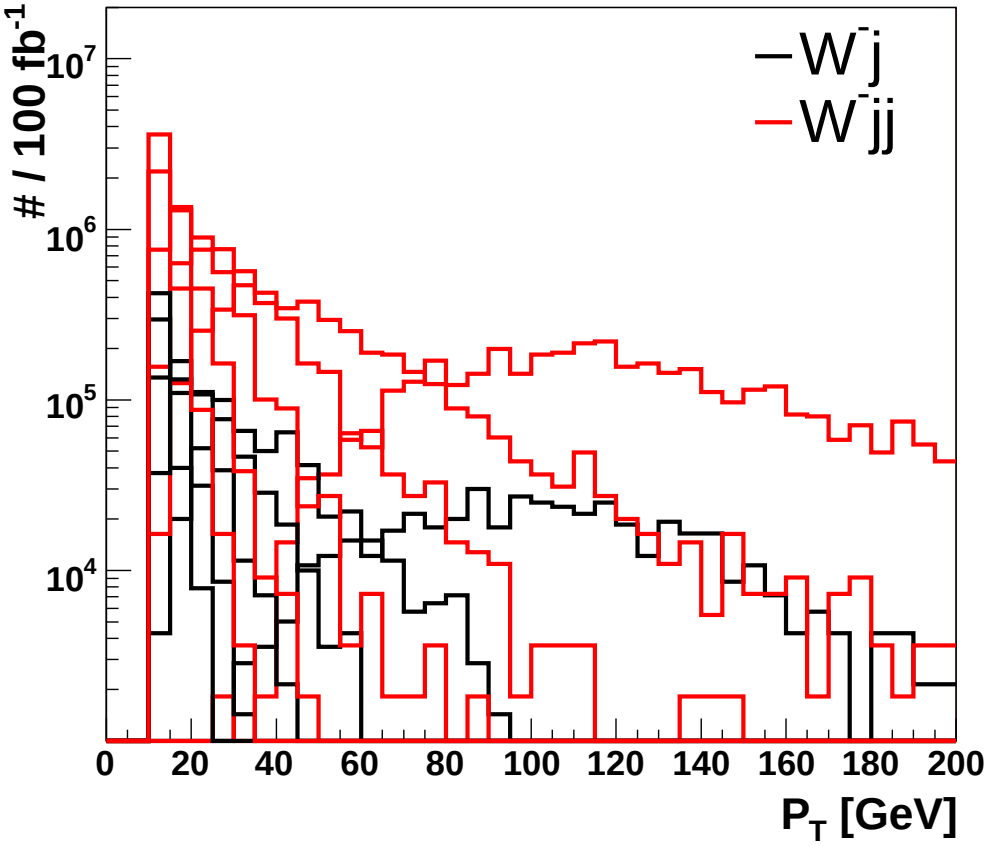


Wjj and Wjjj

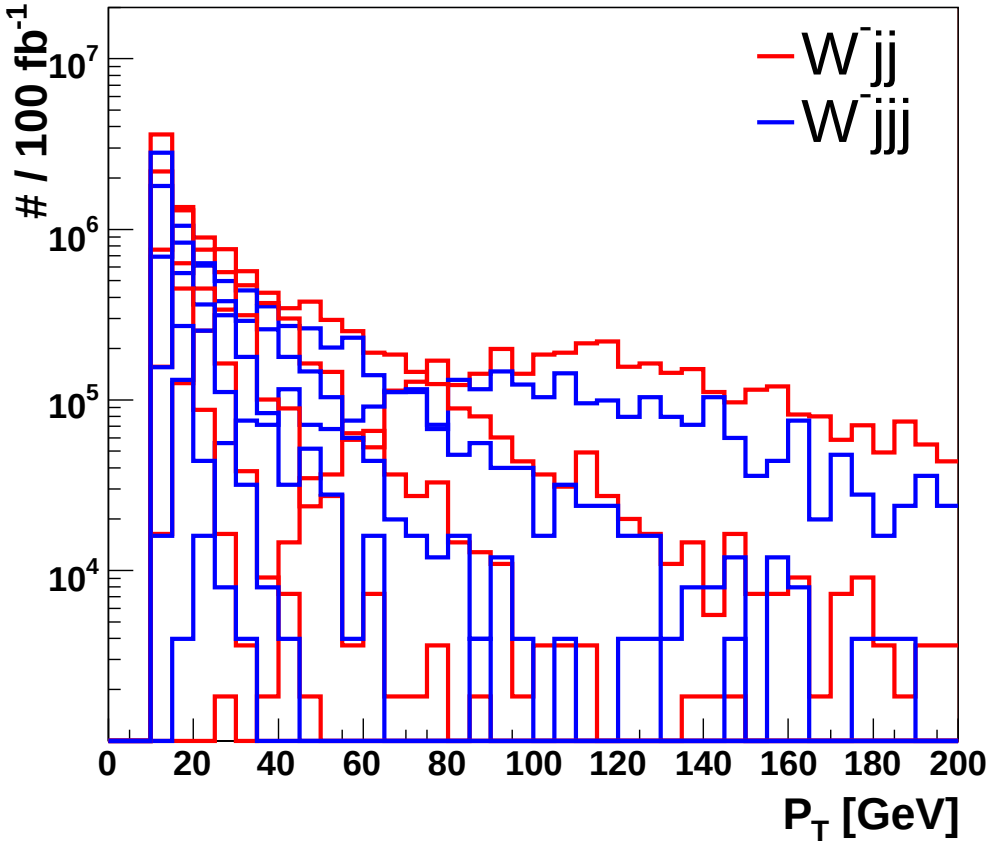


Wm+jets, # clusters=6

Wj and Wjj

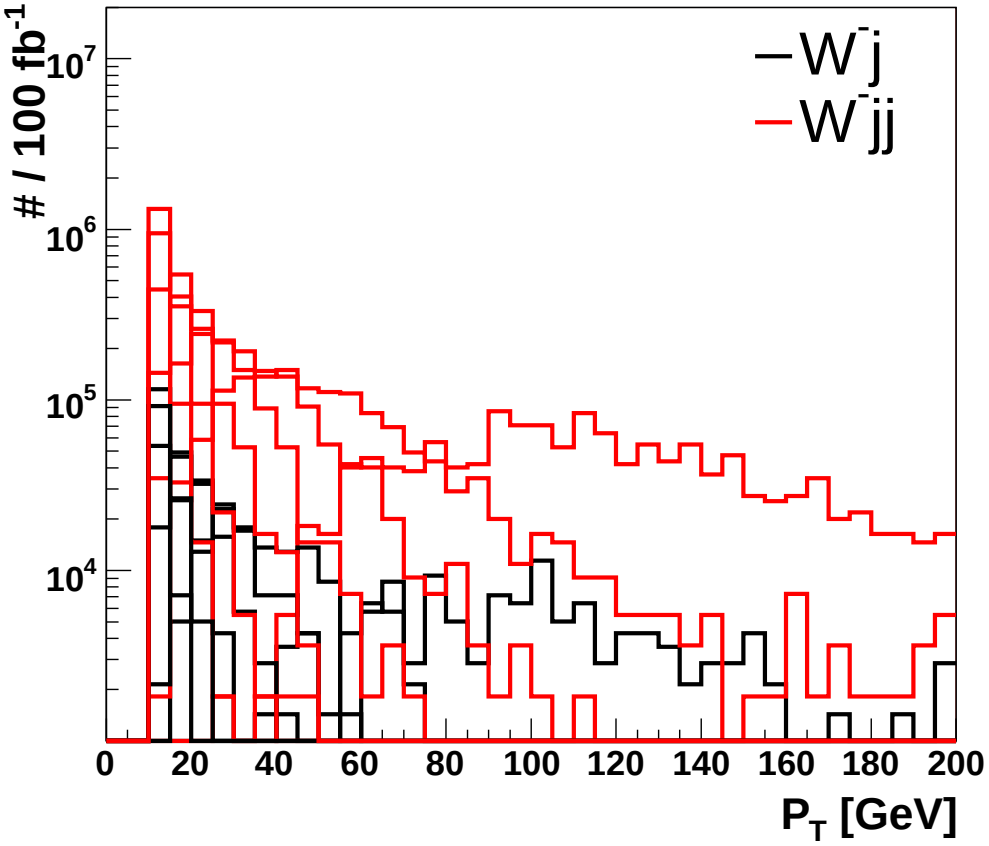


Wjj and Wjjj

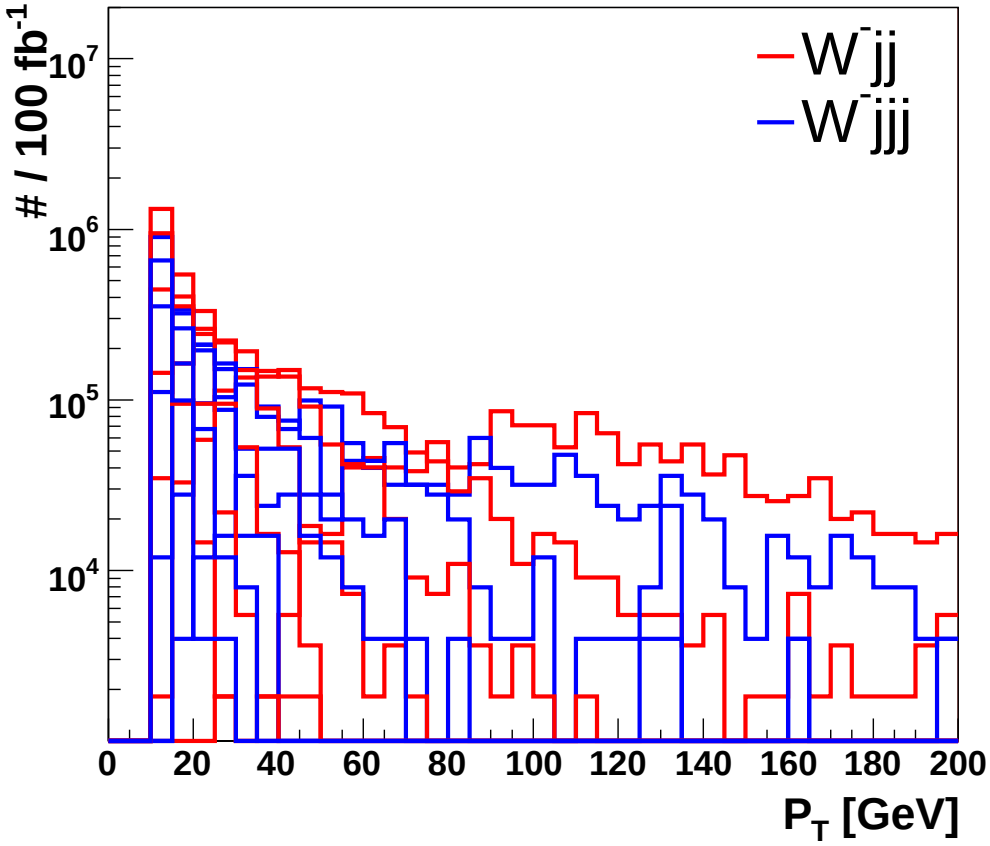


Wm+jets, # clusters=7

Wj and Wjj



Wjj and Wjjj



Summary for W+jets

1. W(jj)W($\mu\nu$)jj signal $\Rightarrow$ not worth to look at # clusters$\geq$5
2. more compicated gener. cuts and more gener. jets $\Rightarrow$ comparison difficult

overlap between Wj and Wjj

- # clusters=0 - not important
- # clusters=1 - a bit problem with normalization, shape the same
- # clusters$\geq$2 - important difference
- for this study: use Wjj

overlap between Wjj and Wjjj

- a bit problem with normalization
- # clusters=0 - not important
- # clusters$\leq$4 - the same
- # clusters=5 - very similar
- # clusters$\geq$6 - important difference appears for !
- for this study: use Wjj

- it is impossible to limit overlap (unless cut on # of clusters)
- recomendation for use $t\bar{t}$ and W with possibly the biggest needed number of jets
- selection results (with det simulation) for $t\bar{t}$ and $t\bar{t}j$:
  $t\bar{t}j$ much more important background (suggestion: j as a tag jet)
- use only one process with defined # of jets to represent $t\bar{t}$+jets and W+jets backgrounds

Conclusions for my study

- use $t\bar{t}j$ and Wjj (or Wjjj)

Additional slides

Selection results for $\mathcal{L}$=100 fb^{-1}