Analysis of the $t\bar{t}$ production in the all-hadronic channel at $\sqrt{s} = 7\,TeV$ with CMS detector:

kinematical selection, mass reconstruction and cross-section measurement

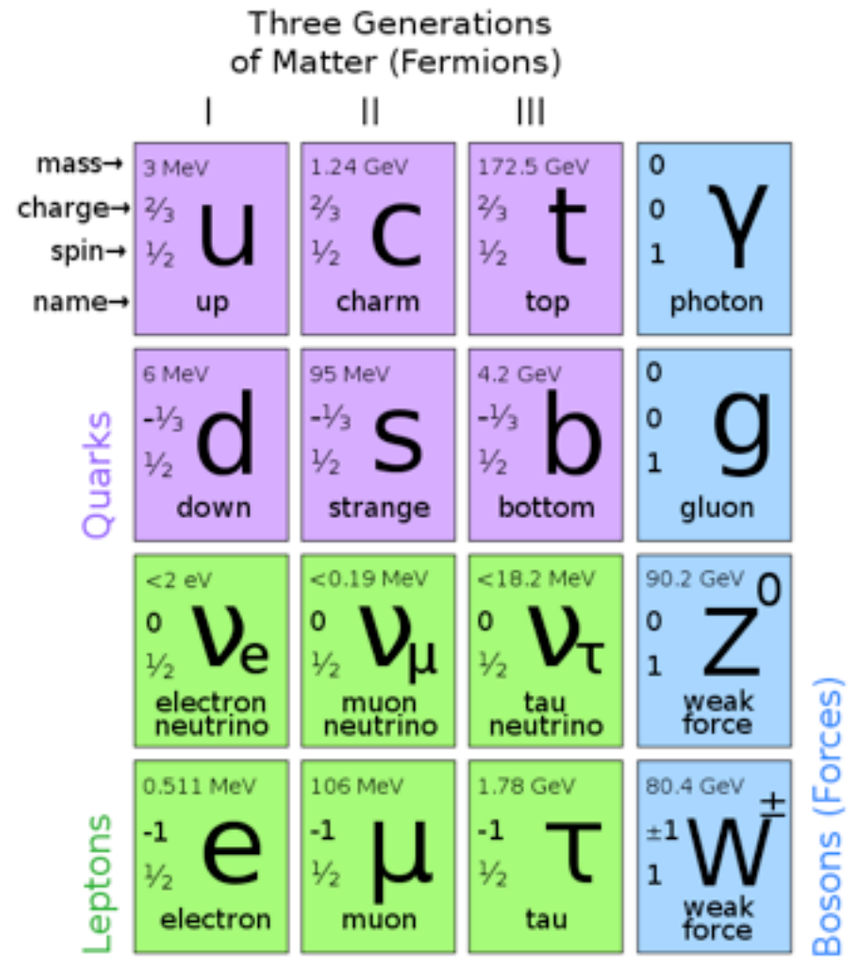
Mauro Sacco

The Top Quark

What is it?

- 3rd generation particle in the SM
- Heaviest particle in the SM

⟶ decays almost immediately into a Wb pair



Motivation of the analysis

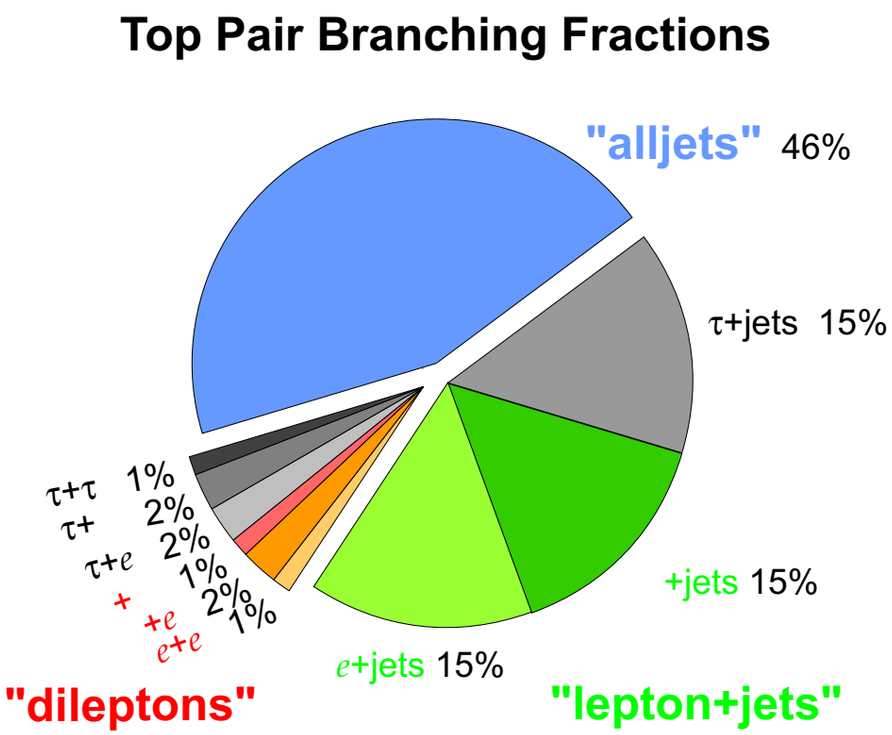
- Measurement of M_t and $\sigma_{t\bar{t}}$
- Check consistency Standard Model predictions
- Test for physics contributions beyond the SM

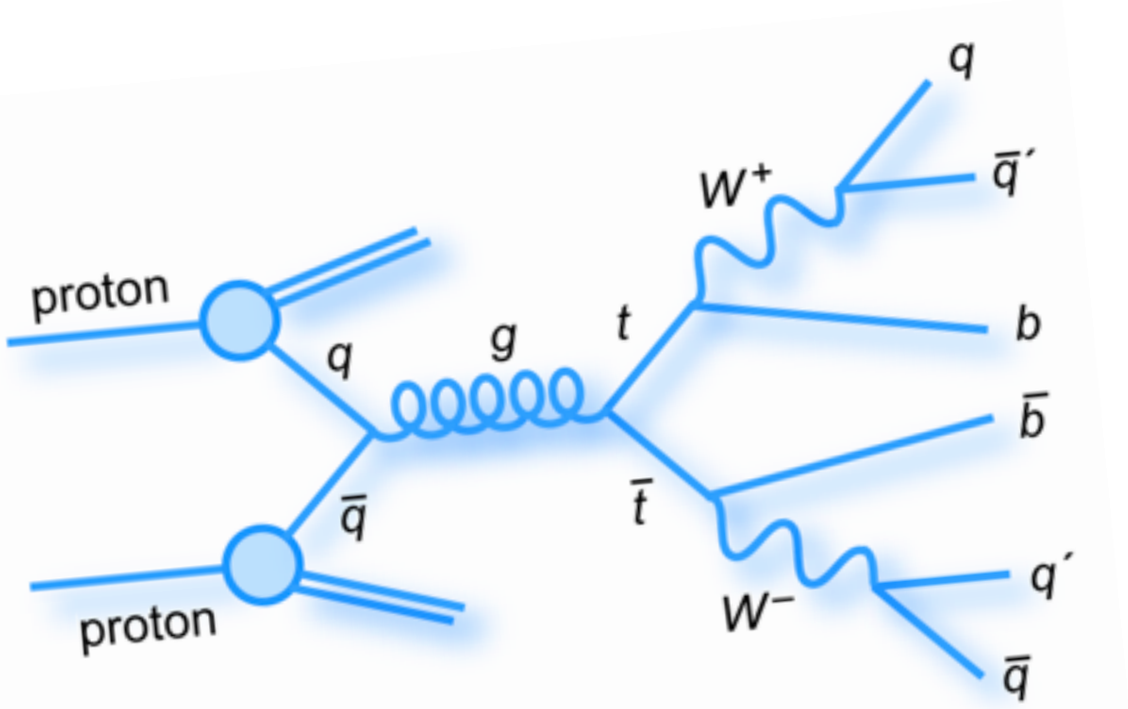
Top Quark Production and Decay

Process: $pp \rightarrow t\bar{t} \rightarrow W^+bW^-b$

Three decay channels:

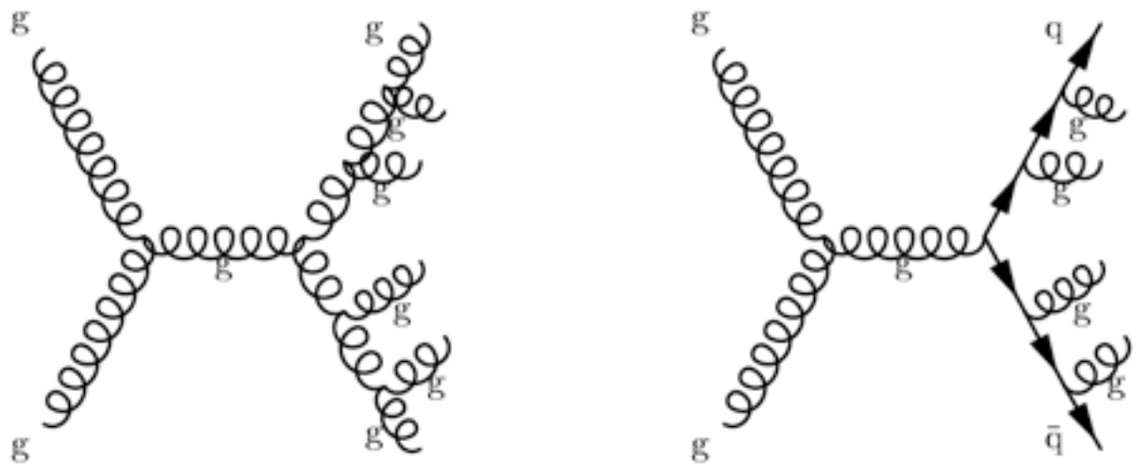
- Dilepton channel
- Single-lepton channel
- **All-hadronic channel**





The all-hadronic channel

- Characterized by the highest branching ratio (46 % of the decays)
- **BUT** overwhelmed by a huge QCD multijet production



dominates by 3 orders of magnitude
(after High Level Trigger)

Event Selection

Datasets:

- May 10, 2011: L=204.2 pb-1 ($S/B \approx 1/600$)
- Signal simulation: MC dataset with 1.16 million events

Isolate the **six-jet decay mode** of top-pair production from the **QCD background**

- First step: Select event with $N_{jet} \geq 6$ + p_T larger than a certain threshold

  Data: 237696 events ($S/B \approx 1/80$)

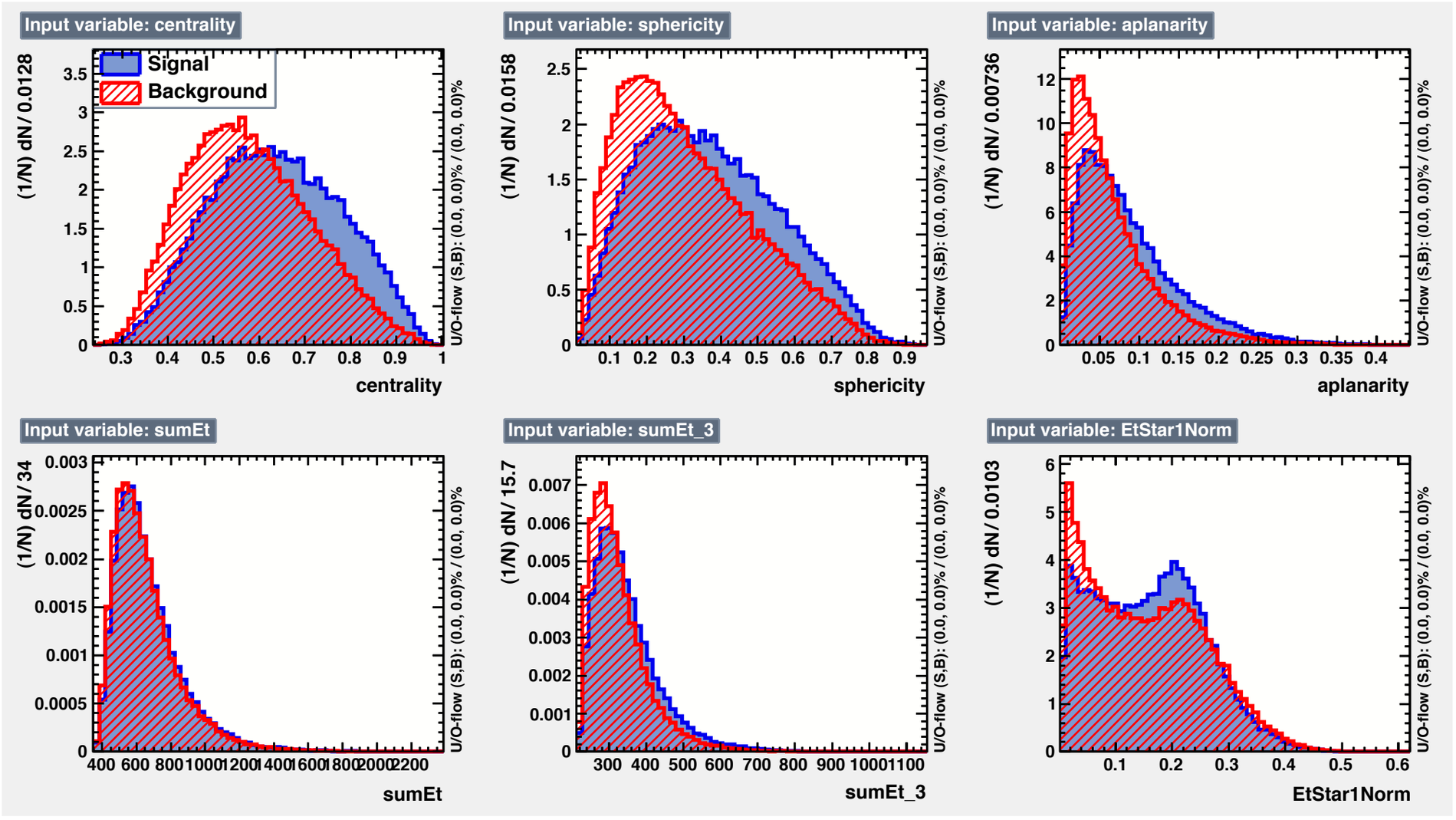
  ⟶

  MC: 120463 events

- Second step: A tight kinematical selection, exploiting a series of variables which have been proven to be effective in the discrimination between signal and background

Kinematical Selection

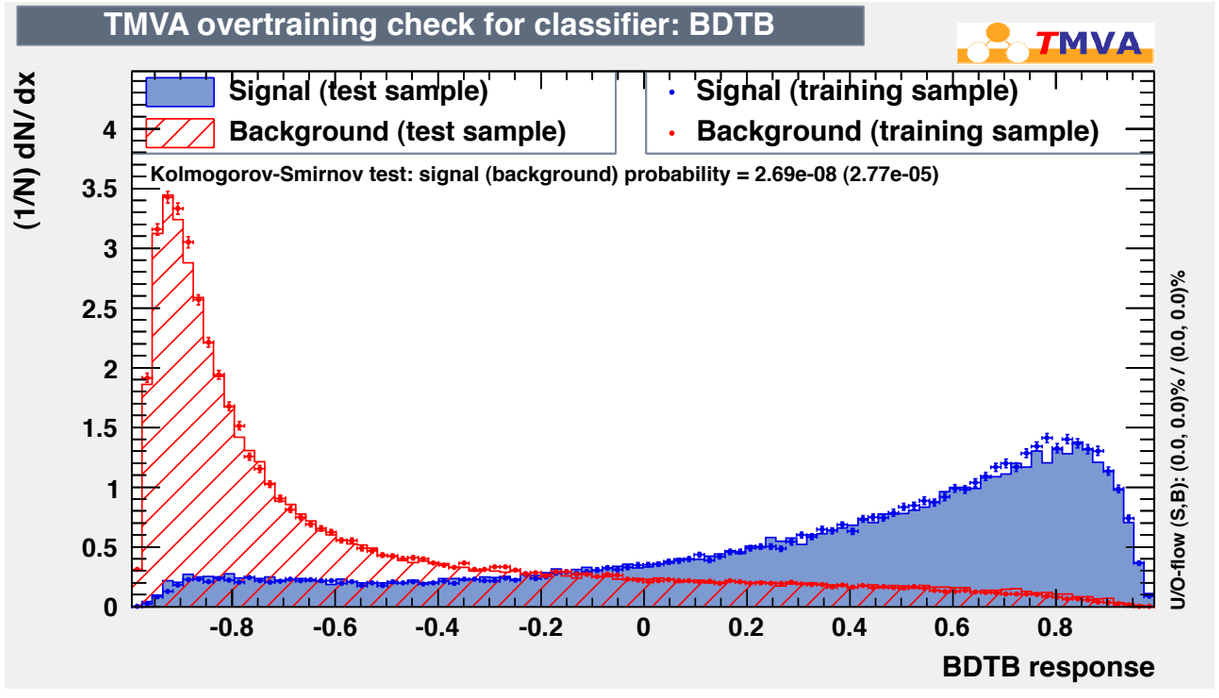
Some discriminative variables: Centrality, Sphericity, Aplanarity, Sum of transverse energy of the jets, . . .



Combine the discriminative power of all the variables:
multivariate analysis

Multivariate Analysis Classification

- Toolkit for Multivariate Analysis integrated in ROOT
- Multivariate techniques: make use of training events, for which the desired output is know (ex. Signal and Background) to determine the mapping function that describes a decision boundary.
- Technique implemented in the analysis:
  Boosted Decision Tree (BDT)



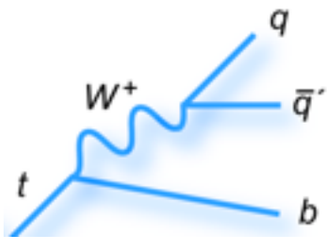
Results

LEVEL	N MC	S	EFFICENCY	B	S/B	S/SQRT(S+B)
$6 \leq N_{jet}$	120462	3340	10.38%	237696	1/81	6.9
Original NN	41809	1155	3.6%	33507	1/28	6.3
BDT1	71046	1970	6.1%	14256	1/6.3	16.5
Original NN + $\geq$ 2 Medium b-tag	20958	579	1.8%	3337	1/5	10.0
BDT1 + $\geq$ 2 Medium HE b-tag	47413	1315	4.1%	8057	1/5	14.7

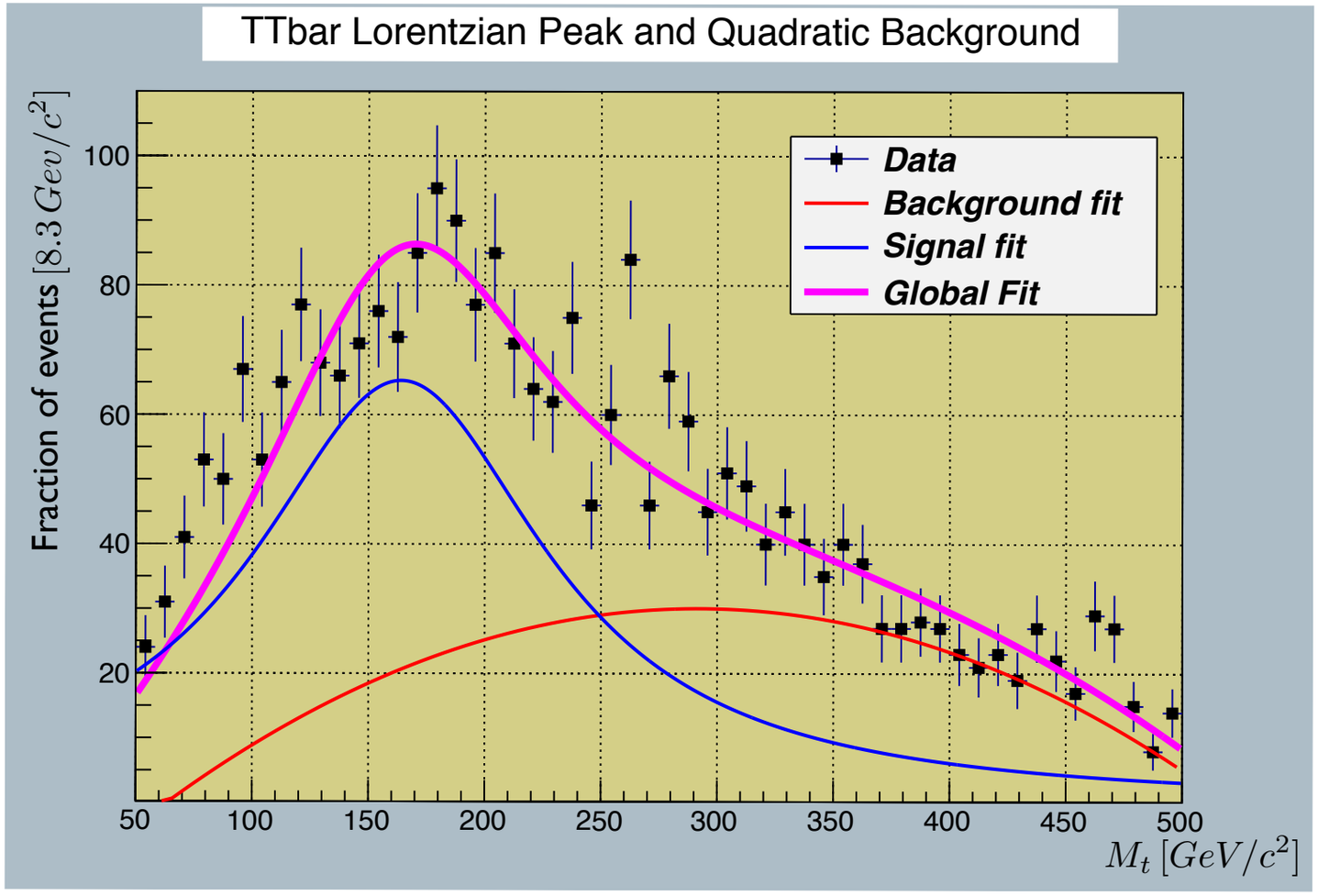
- We use b-tagging in order to reach a better S/B
- Comparing the BDT and the Neural Network approaches we see that BDT > NN !
- BDT + b-tagging give a good S/B!
- Apply the same procedure to Summer 2011 Dataset (L=1096 pb-1) -> S/B=1/2 !!!!

→ Compute cross section ←

Mass distribution and cross-section



→ Invariant mass of top



Cross section: $\sigma_{t\bar{t}} = 203 \pm 35(stat.) \pm 12(lumi.)\ pb$

Theoretical: $\sigma_{t\bar{t}} = 165 \pm 10\ pb$