Boosted tops in the ttbar dilepton channel: optimization of the lepton selection

Ibles Olcina Samblàs
Supervisor: Carmen Diez Pardos

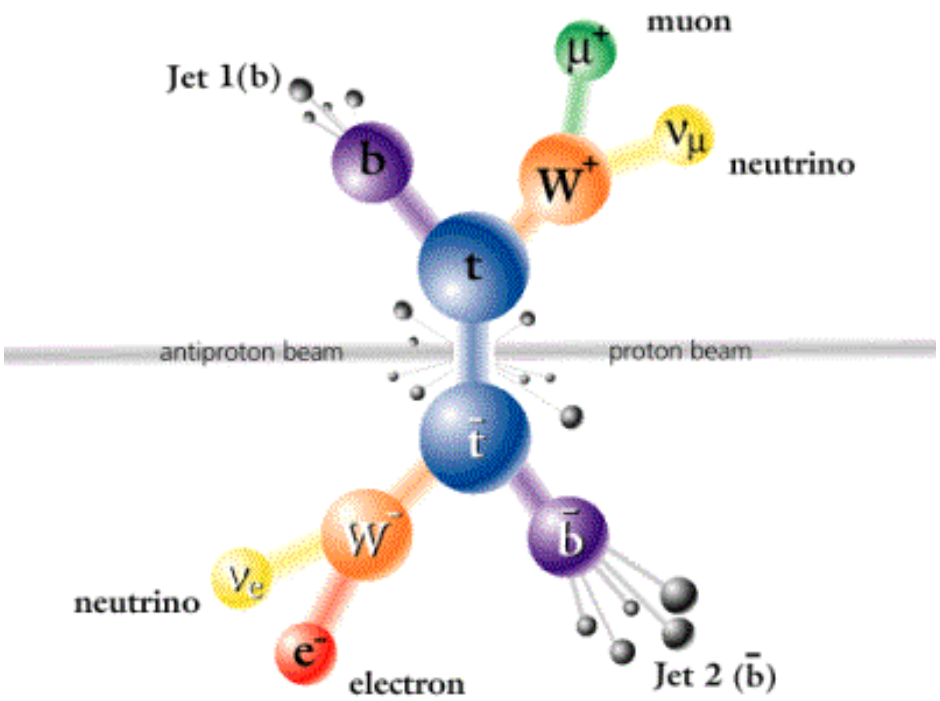
DESY Summerschool 2014

Top physics group CMS-DESY

Table of Contents

1. Introduction
2. Kinematics of boosted topologies
3. Comparison between two different isolation criteria
   - Regular isolation criterion
   - 2D cut
4. Results with the optimized selection
5. Summary

The ttbar dilepton channel



Schematic of the production of a $t\bar{t}$ pair in the dilepton channel and one of its possible decays

Dilepton channel

- $t \rightarrow b\ l^{+}\nu_{l}$
- $\bar{t} \rightarrow \bar{b}\ l^{-}\ \bar{\nu}_{l}$

Low BR ($\sim 5\%$) but also very low background

Data sample used

- LHC 2012 run period
- pp collisions at $\sqrt{s} = 8$ TeV and an integrated luminosity of $19.7 \pm 0.9\ fb^{-1}$

Motivation

- Study the effects of having top quarks with very high momentum
- Optimise the selection of lepton events for this case

Event selection: requirements

Events selected using a trigger based on dileptons

2 isolated, opposite-charged leptons

Leptons with high momentum:

- $p_T^l > 20$ GeV within $|\eta| < 2.4$

2 energetic jets

- At least 2 jets with $p_T > 30$ GeV within $|\eta| < 2.4$
- After applying a b-tagging algorithm: At least 1 b-tagged jet

Additional requirements to reject background

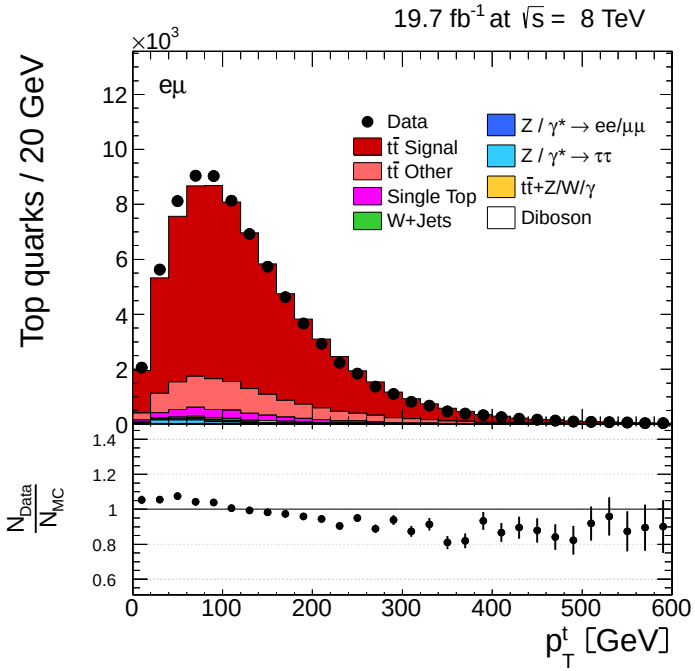
Reject Z+jet background for the $\mu^+\mu^-$ and e^+e^- channels:

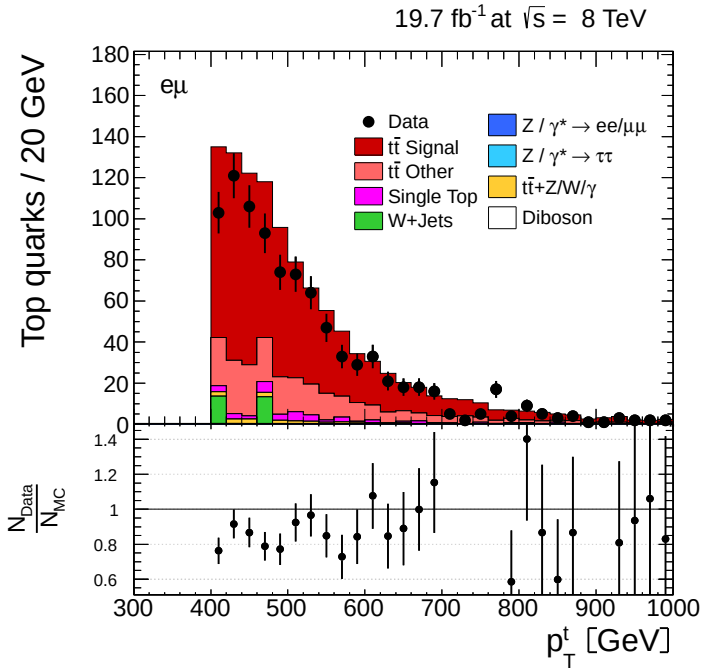
- $\not{E}_T > 40$ GeV (**Def:** $\not{E}_T = -\sum_{lep} p_T - \sum_{jets} p_T$)
- Exclude Z boson mass region: 76 GeV $< m^{ll} <$ 106 GeV.

Kinematic reconstruction

Kinematics of the full $t\bar{t}$ event are totally reconstructed solving the kinematic equations.

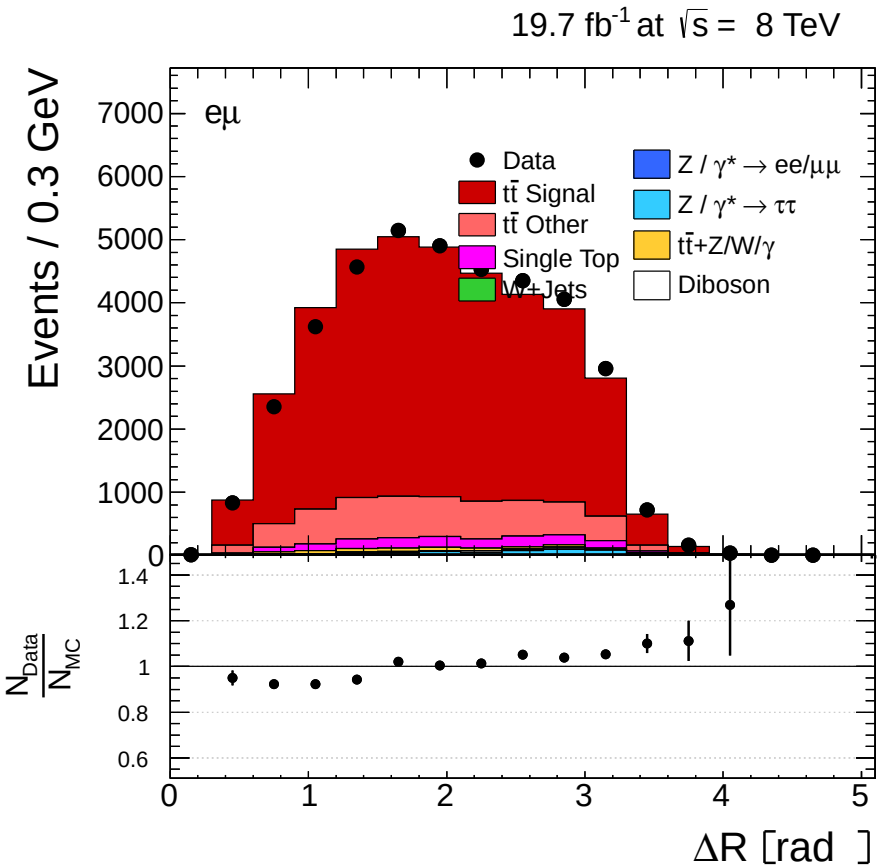
Boosted tops: $p_T^t > 400$ GeV



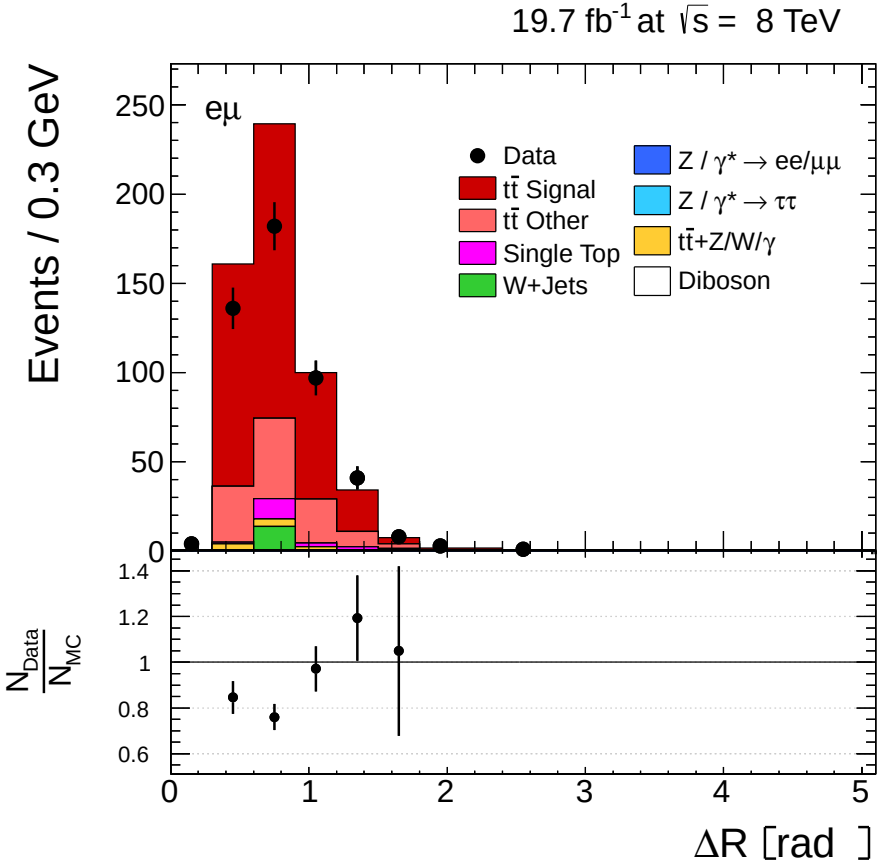


	Full data set		Boosted regime	
	Events	Ratio over MC (%)	Events	Ratio over MC (%)
Total MC	38396		548	
Total Data	38239	99.6	472	86.2
Total Background	7708	20.1	157	28.7
$t\bar{t}$ signal	30688	79.9	391	71.3

$\Delta R\left(l, \bar{b} - jet\right)$ where $\Delta R = \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2}$

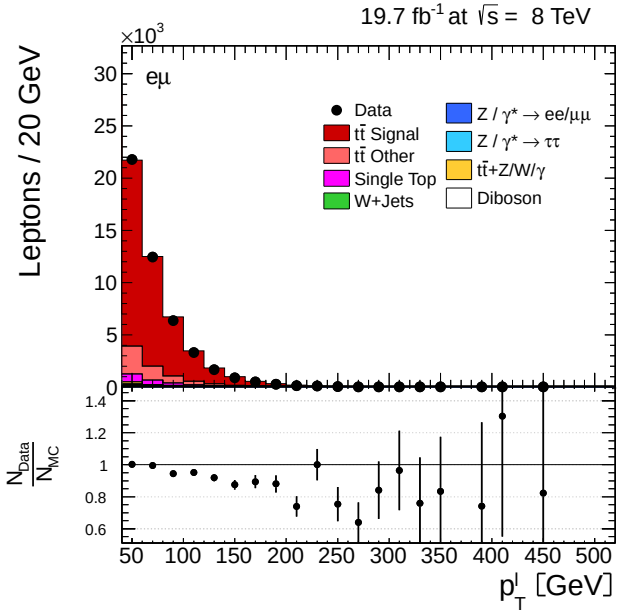


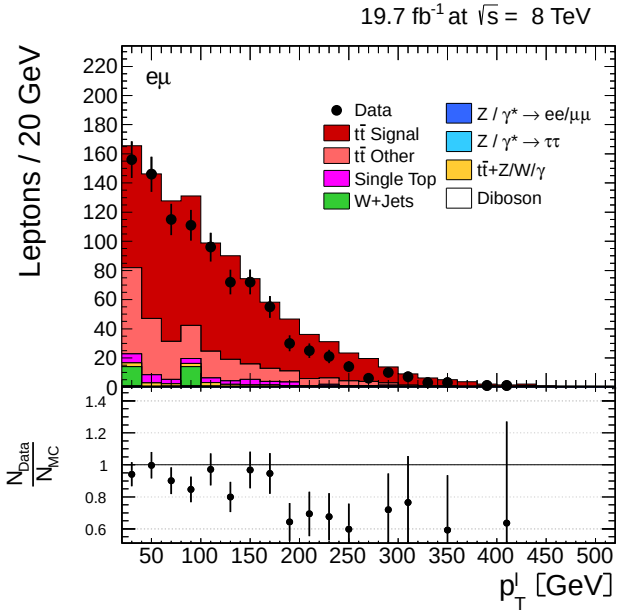
Full data set



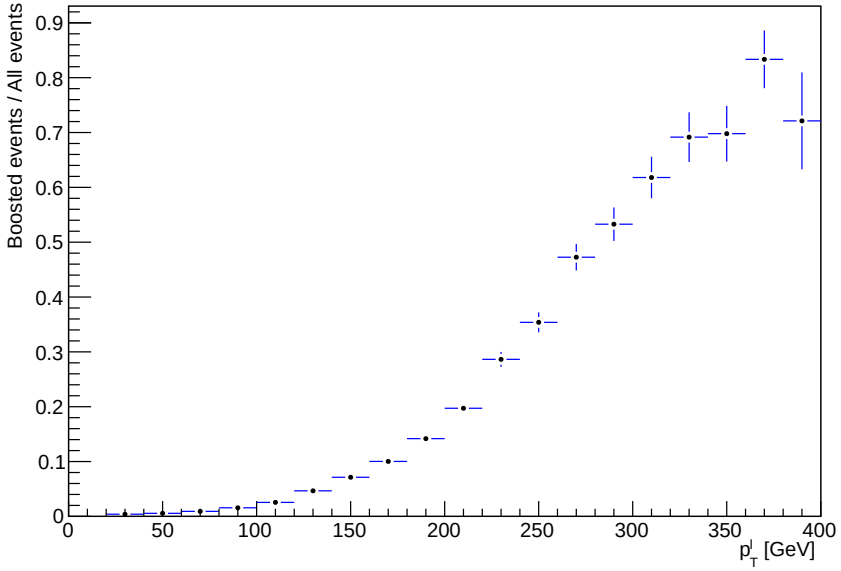
Boosted regime

Transverse momentum distribution of the leptons





There is a **very large fraction** of leptons with high momentum coming from boosted tops:



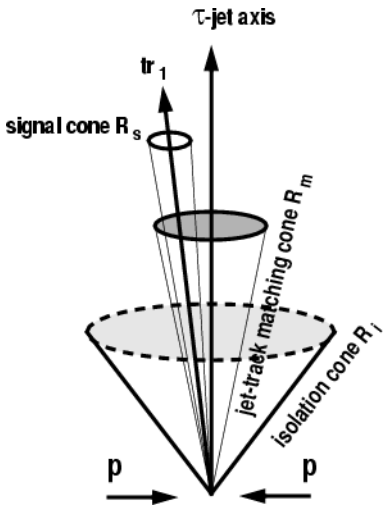
First optimization: Change the cut to the momentum of leptons from 20 to 40 GeV
Second optimization: Leptons and jets are closer in boosted topologies → check isolation efficiency

Regular isolation criterion

Definition

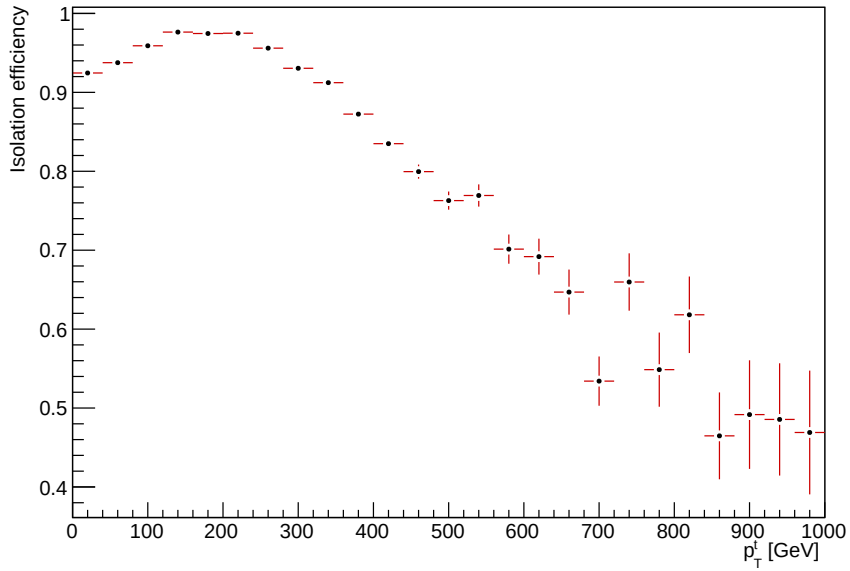
- $I^l_{rel} < 0.15$ inside a cone of $\Delta R < 0.3$, where $I^l_{rel} = \sum_{i=had,\gamma} E^i_T / p^l_T$

Meaning: Most of the energy inside a cone of $\Delta R < 0.3$ around the lepton comes from the lepton.



Isolation efficiency

Events **with** isolation condition / Events **without** isolation condition



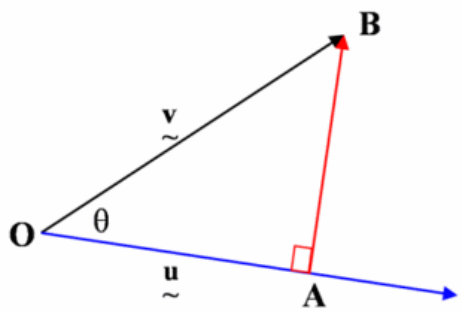
A **large fraction** of top events with high momentum are lost because of this isolation condition

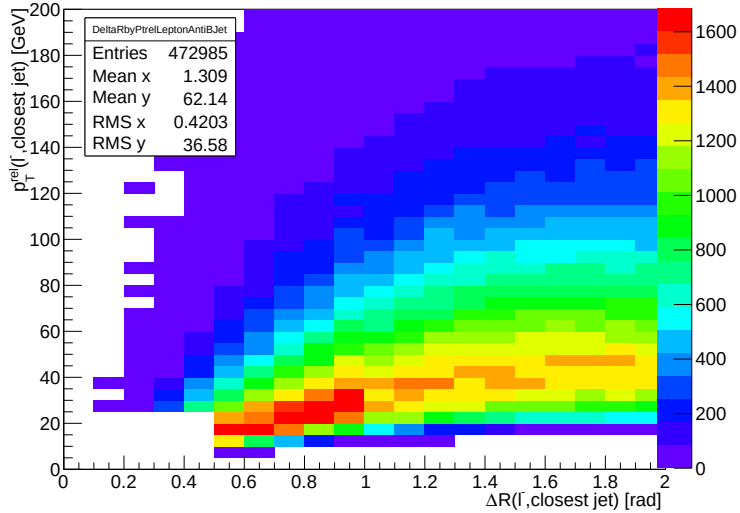
2D cut

Definition

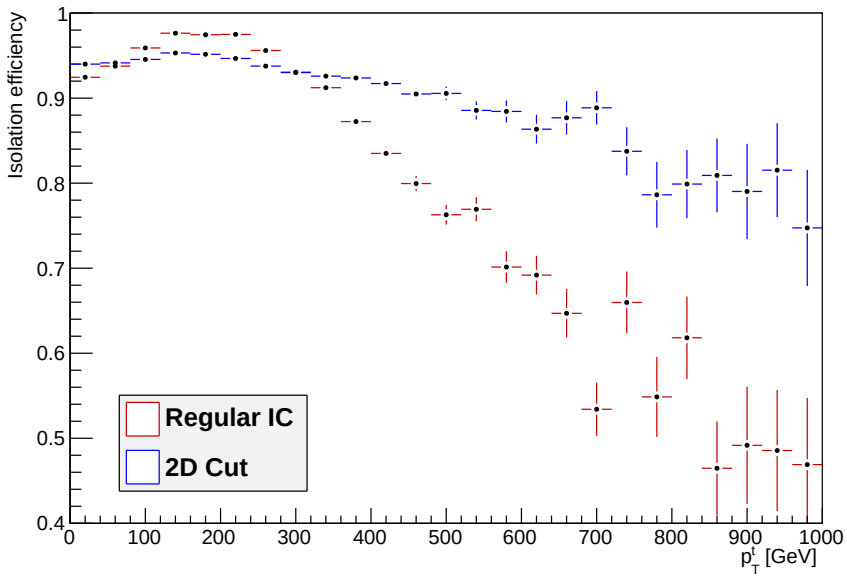
- $\Delta R(l, closest\ jet) > 0.5$ OR
- $|p_T^{rel}(l, closest\ jet)| > 25$ GeV

Meaning: Reject event if the lepton is too close to its closest jet.

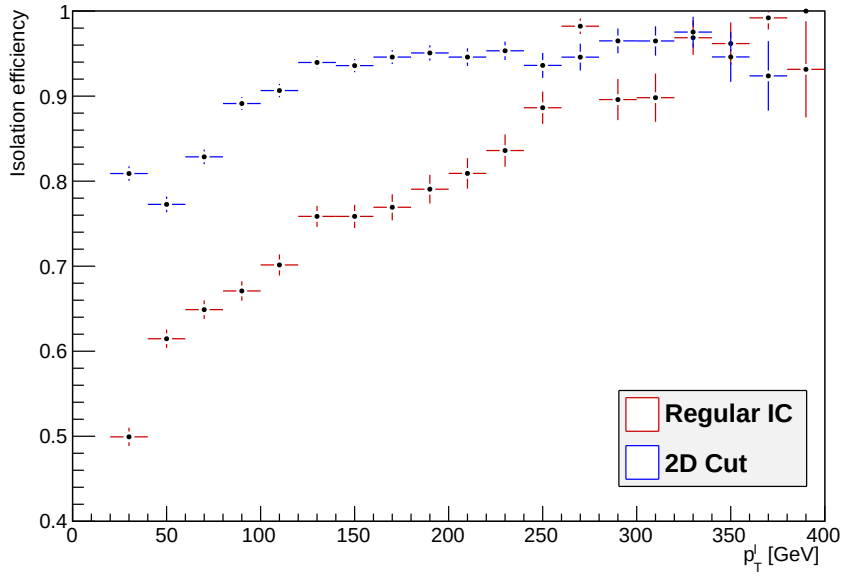




Comparison of isolation efficiencies



Top quark transverse momentum



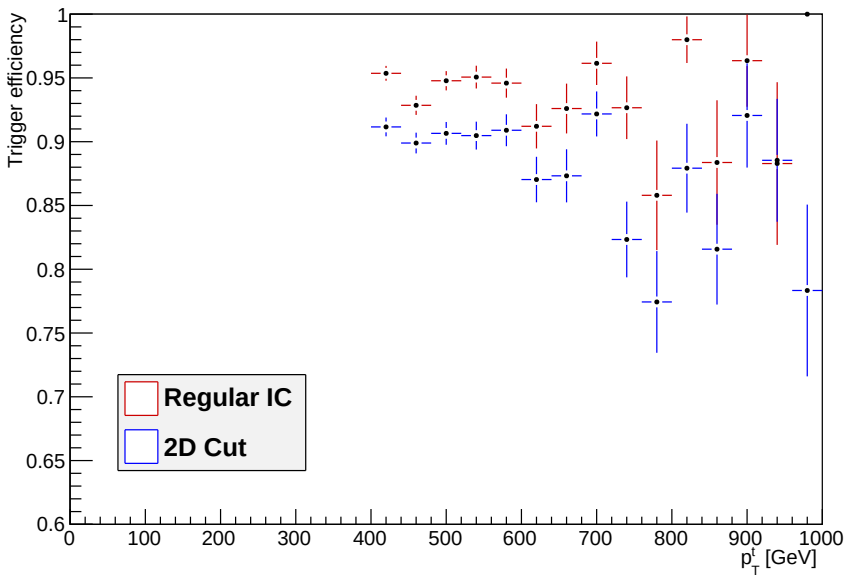
Transverse momentum of leptons coming from boosted tops

	No Isolation	Regular Isolation Cond.	2D cut
Full data set	469277	444696	442814
Loss of events (%)		5.2	5.6
Boosted regime	7651	5595	6708
Loss of events (%)		26.9	12.3

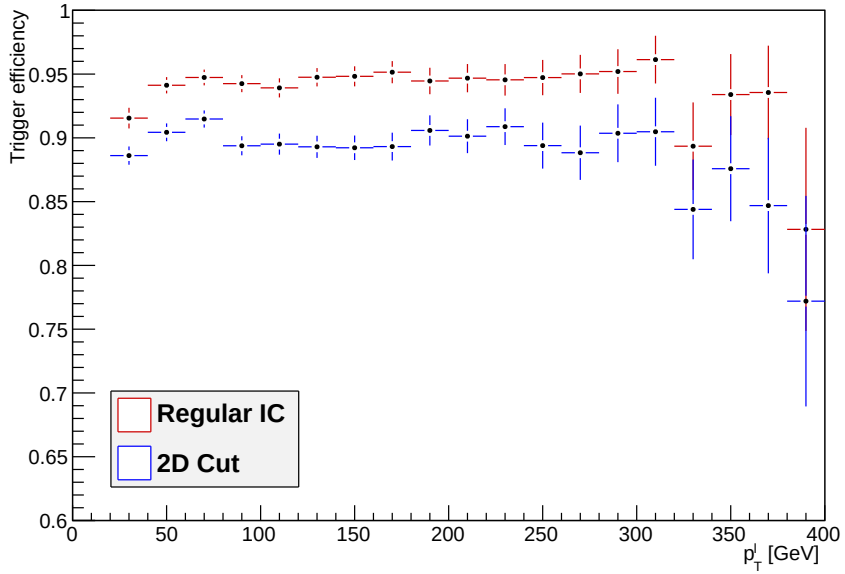
Comparison of trigger efficiencies

Trigger efficiency

Events **with** trigger condition / Events **without** trigger condition

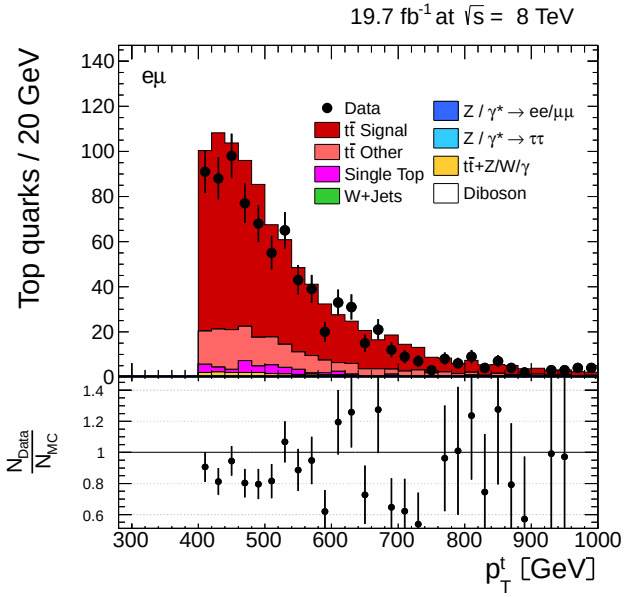


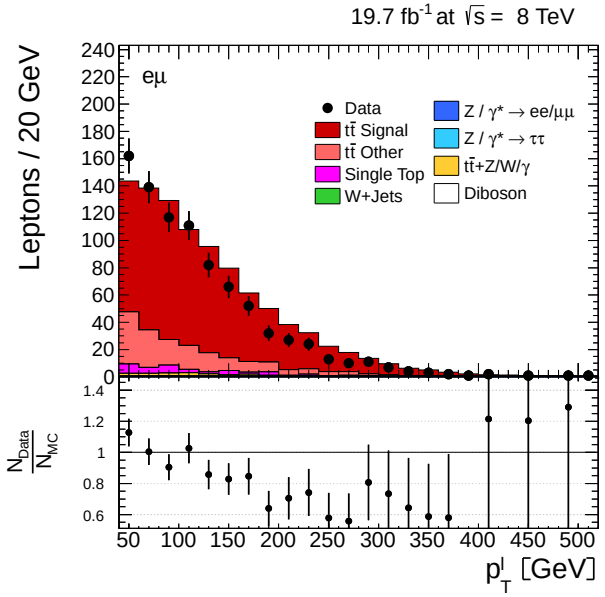
Top quark transverse momentum

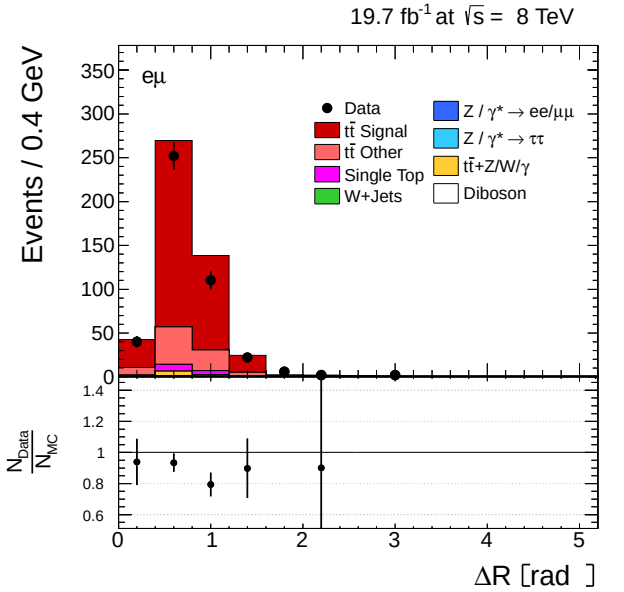


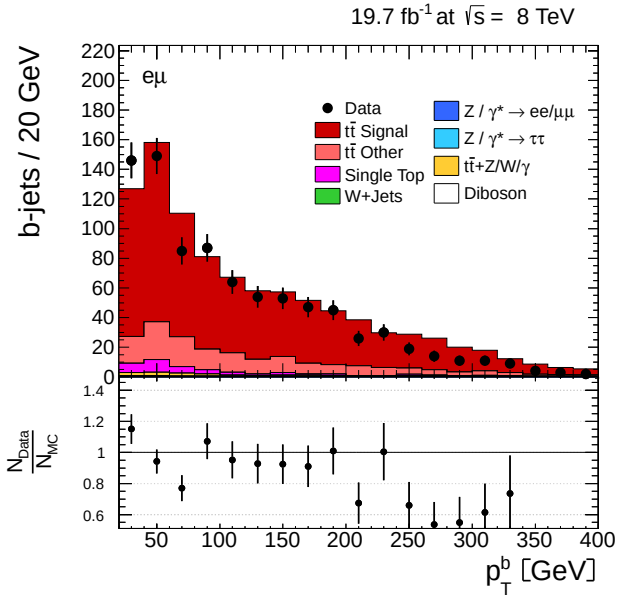
Lepton transverse momentum of leptons coming from boosted tops

Boosted sample









Monte Carlo and data yields ($e\mu$ channel)

Regular IC	**Boosted regime**	
	Events	Ratio over MC (%)
Total MC	548	
Total Data	472	86.2
Total Background	157	28.7
$t\bar{t}$ signal	391	71.3

2D cut	**Boosted regime (20 GeV cut)**		**Boosted regime (40 GeV cut)**	
	Events	Ratio over MC (%)	Events	Ratio over MC (%)
Total MC	672		482	
Total Data	742	110.4	434	90.0
Total Background	204	30.3	106	22.0
$t\bar{t}$ signal	468	69.7	376	78.0

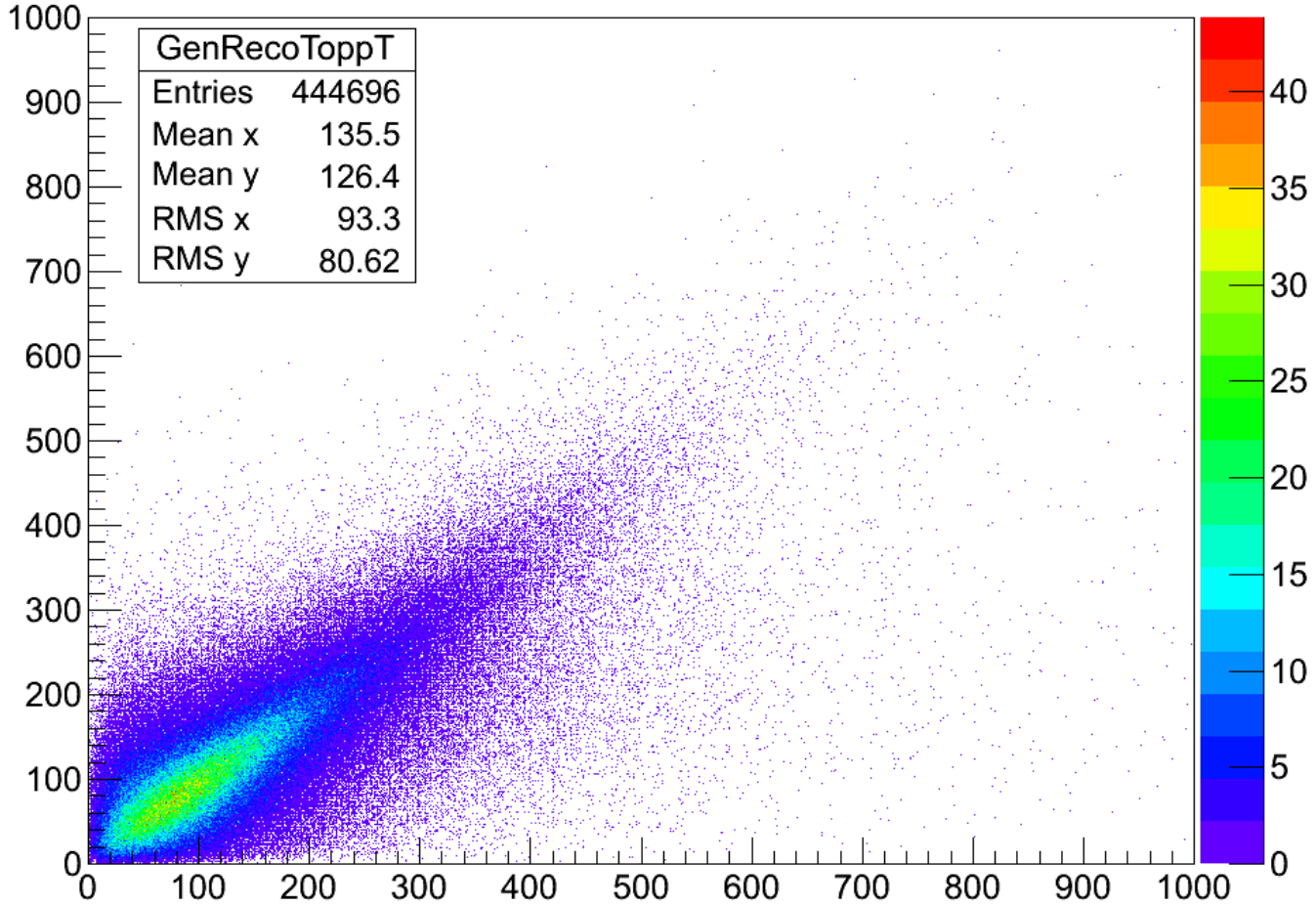
Summary

- Boosted regime: exploration of a different kinematic region
  - $\Delta R(l, \bar{b} - jet)$ decreases and p_T distribution of leptons is shifted towards higher momenta.
  - There is a small increase of the background and the agreement between MC and data gets worse.
- Standard selection of leptons needs optimization
  - Increase the p_T cut from 20 GeV to 40 GeV
  - Implement "2D cut" isolation criterion, which works better than the regular isolation criterion:
    - ⋆ Approximately 15% improvement of the isolation efficiency
    - ⋆ The trigger efficiency is only about 4% worse
  - With these modifications there is a better agreement MC-data and the proportion of background is reduced considerably for boosted topologies.
- Further studies:
  - Look for a possible optimization of the kinematic reconstruction too
  - Carry out a similar study for $t\bar{t}H$ processes, which have a similar signature


THANK
YOU

BACKUP

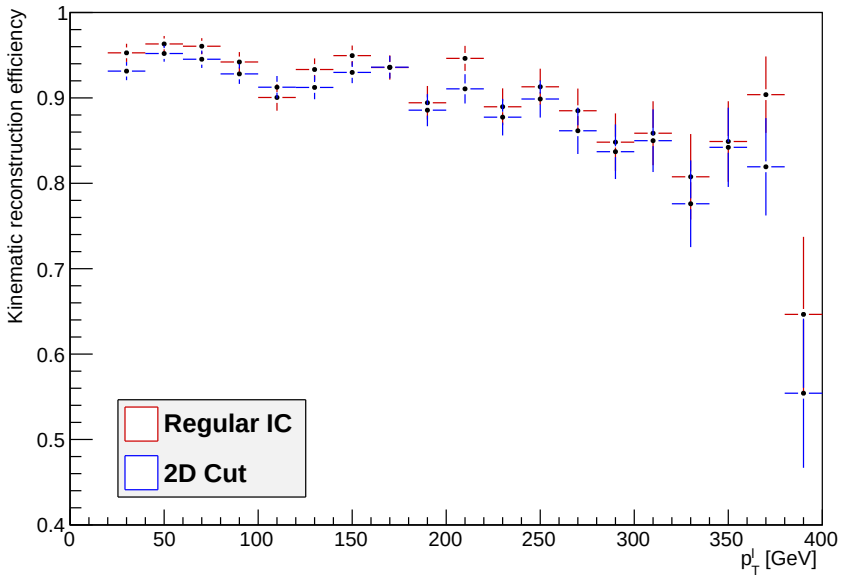
Kinematic reconstruction efficiency



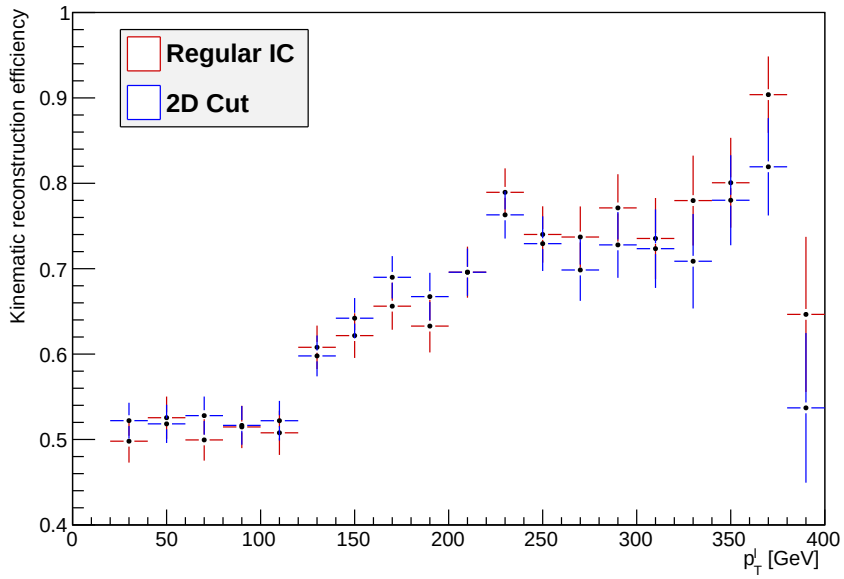
$x-$axis: Transverse momentum of the reconstructed top quarks

$y-$axis: Transverse momentum of the generated top quarks

Kinematic reconstruction efficiency: Lepton p_T

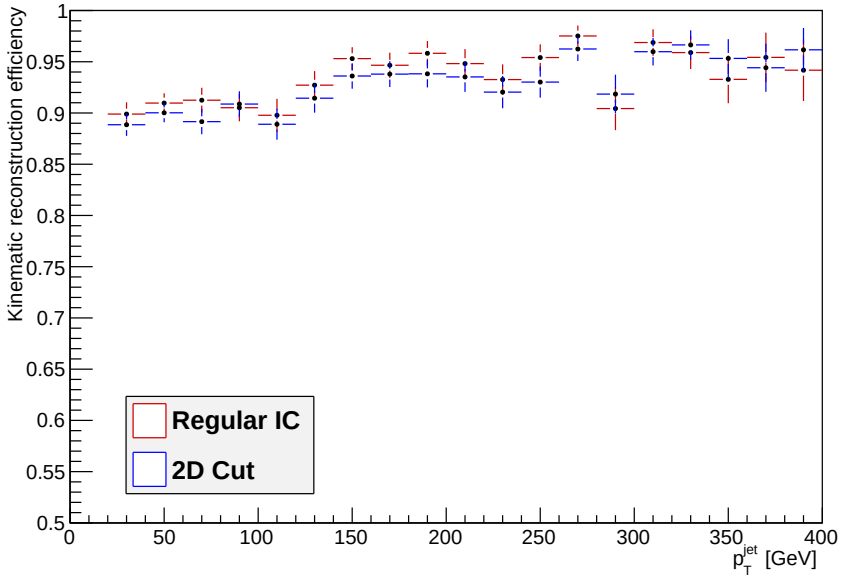


Boosted condition only applied to the generated tops

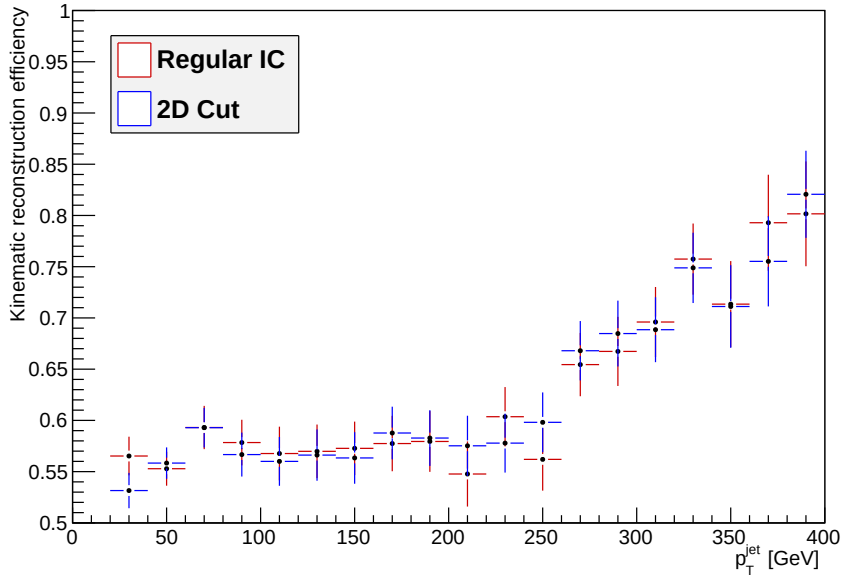


Boosted condition applied to both the generated and reconstructed tops

Kinematic reconstruction efficiency: Jets p_T

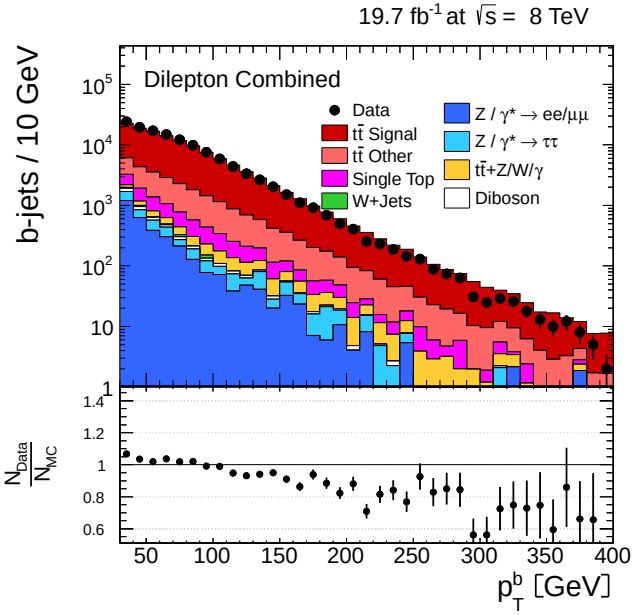


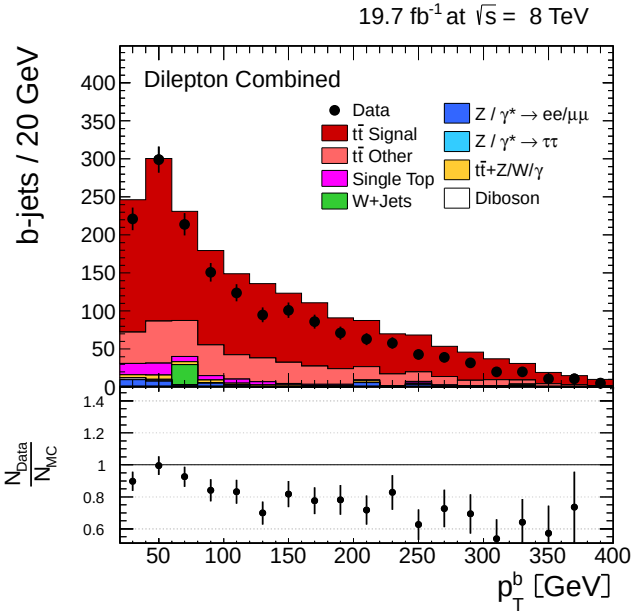
Boosted condition only applied to the generated tops



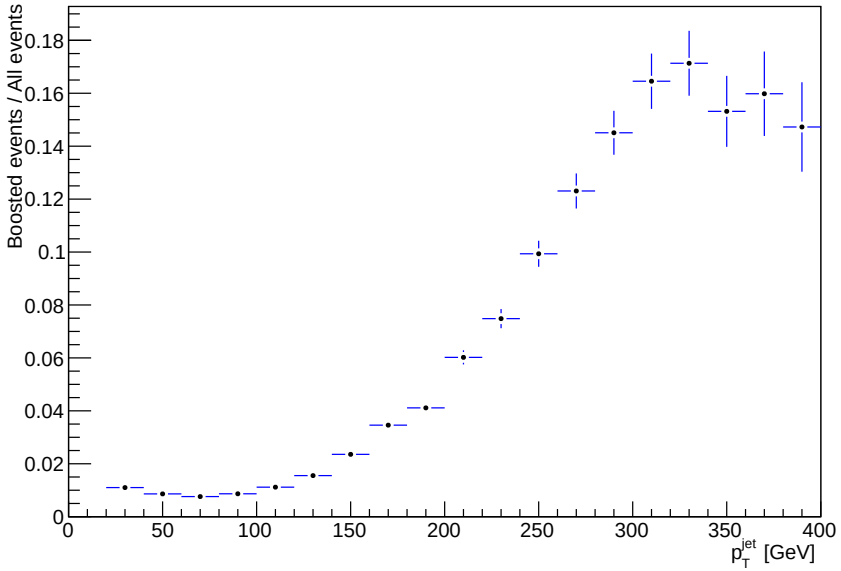
Boosted condition applied to both the generated and reconstructed tops

Transverse momentum distribution of the b-jets





Ratio of boosted events over all the events for the $t\bar{t}$ signal



There is a **moderate fraction** of b-jets with high momentum coming from boosted tops.