HEP Top Tagger for the search of stops

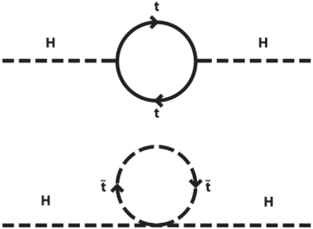
Giulio Dujany

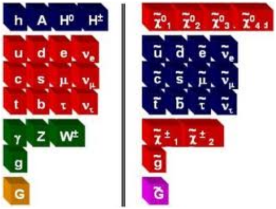
Summer intern 2012
Supervised by: Daryl Hare,
Rick Cavanaugh, Chris Silkworth

26 September 2012

Stop

SUSY can solve the Hierarchy problem





Stop should be the lightest squark

- Top is the heaviest fermion: his loop contribute the most to Higgs' mass correction.
- In MSSM the running masses of squarks are degenerate at high scale.
- Due to large Yukawa coupling stop mass run down at lower energies.

$$pp \to \tilde{t}\tilde{t}^* \to (t\chi_1^0)(\bar{t}\chi_1^0) \to (bjj\chi_1^0)(\bar{b}jj\chi_1^0)$$

Fully Hadronic Decay Channel

Advantages

- Fully reconstructs top kinematics (no $\not{E}_T$ from neutrinos)
- $\not{E}_T$ due only to neutralinos
- Higher branching ratio (color factor)

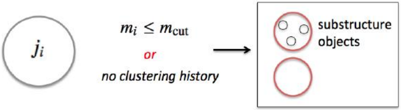
Challenges

Needs an efficient identification of boosted tops

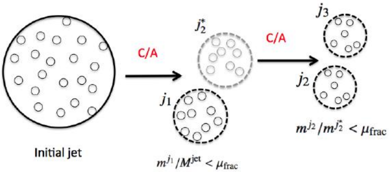
$$\Downarrow$$

We expect a moderate stop mass $\Longrightarrow$ Rate for highly boosted tops in the final state is significantly lower than the overall signal rate $\Longrightarrow$ We need a top-tagger for moderately boosted tops

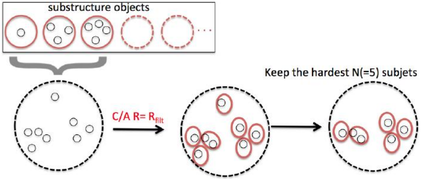
HEP Top Tagger: Top Candidate



(a) Every object encountered in the de-clustering process is considered a ‘substructure object’ if it is of sufficiently low mass or has no clustering history.

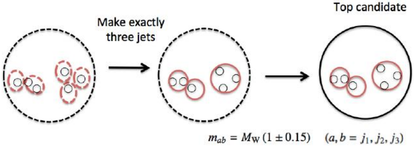


(b) The mass-drop criterion is applied iteratively, following the highest subjet-mass line through the clustering history, resulting in N_i substructure objects.



$$R_{\mathrm{filt}} = \min[0.3, \frac{\Delta R_{j_1,j_2}}{2}, \frac{\Delta R_{j_1,j_3}}{2}, \frac{\Delta R_{j_2,j_3}}{2}]$$

(c) For every triplet-wise combination of the substructure objects, recluster into subjets and select the N_{subjet} leading-p_{T} subjets, with $3 \leq N_{\mathrm{subjet}} \leq N_i$ (here, $N_{\mathrm{subjet}} = 5$).



(d) Recluster the constituents of the N_{subjet} subjets into exactly three subjets to make the top candidate for this triplet-wise combination of substructure objects.

HEP Top Tagger: Selection requirements

pt cut: combined $p_T > 200$ GeV.

top mass cut: $m_{top} - 25\text{ GeV} < m_{123} < m_{top} + 25\text{ GeV}$.

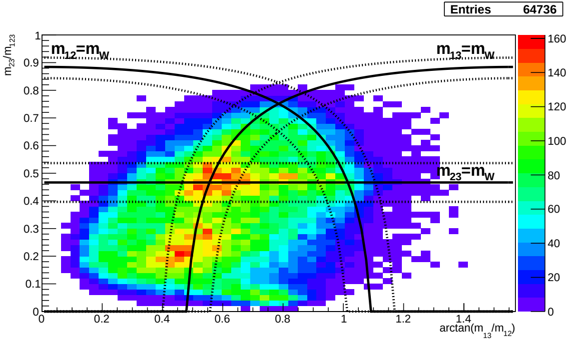
di-mass cut: One of the three:

- $0.2 < \arctan\frac{m_{13}}{m_{12}} < 1.3$ and $R_{min} < \frac{m_{23}}{m_{123}} < R_{max}$
- $R_{min}^2\left(1+\left(\frac{m_{13}}{m_{12}}\right)^2\right) < 1-\left(\frac{m_{23}}{m_{123}}\right)^2 < R_{max}^2\left(1+\left(\frac{m_{13}}{m_{12}}\right)^2\right)$ and $\frac{m_{23}}{m_{123}} > 0.35$
- $R_{min}^2\left(1+\left(\frac{m_{12}}{m_{13}}\right)^2\right) < 1-\left(\frac{m_{23}}{m_{123}}\right)^2 < R_{max}^2\left(1+\left(\frac{m_{12}}{m_{13}}\right)^2\right)$ and $\frac{m_{23}}{m_{123}} > 0.35$

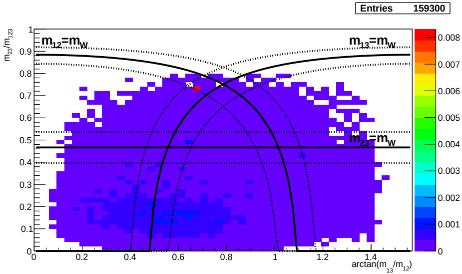
where $R_{min} = 85\% m_W/m_{top}$ and $R_{max} = 115\% m_W/m_{top}$.

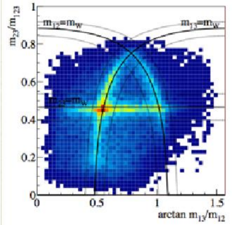
Di-mass plots

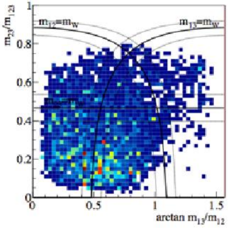
$t\bar{t}$



QCD

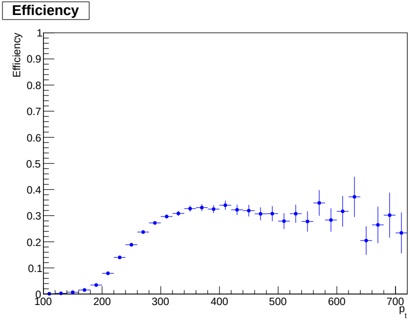






Plehn, Spannowsky, Takeuchi and Zerwas arXiv 1006.2833

Efficiency and fake rate

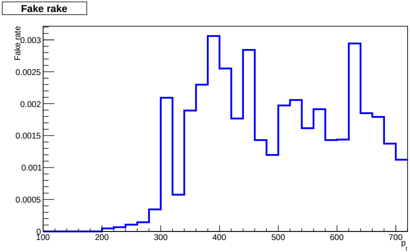


$$\text{Efficiency} = \frac{\#\ \text{tagged tops}}{\#\ \text{generated tops}}$$

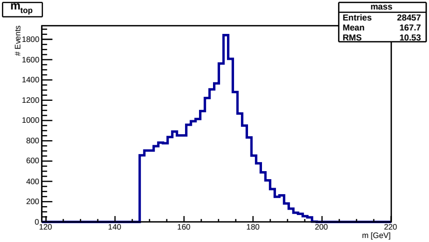
From a $t\bar{t}$ Monte Carlo sample with 450k events.

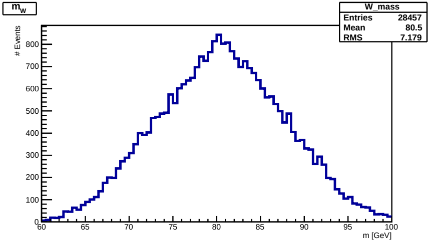
$$\text{Fake rate} = \frac{\#\ \text{mistagged jets}}{\#\ \text{ac1.5 jets}}$$

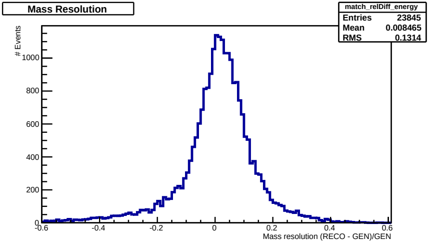
From a weighted QCD Monte Carlo sample with 750k events.

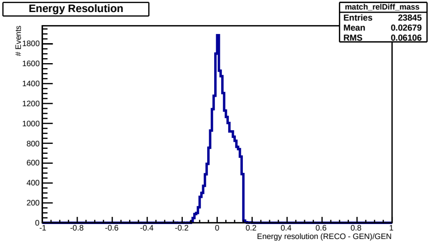


Performance

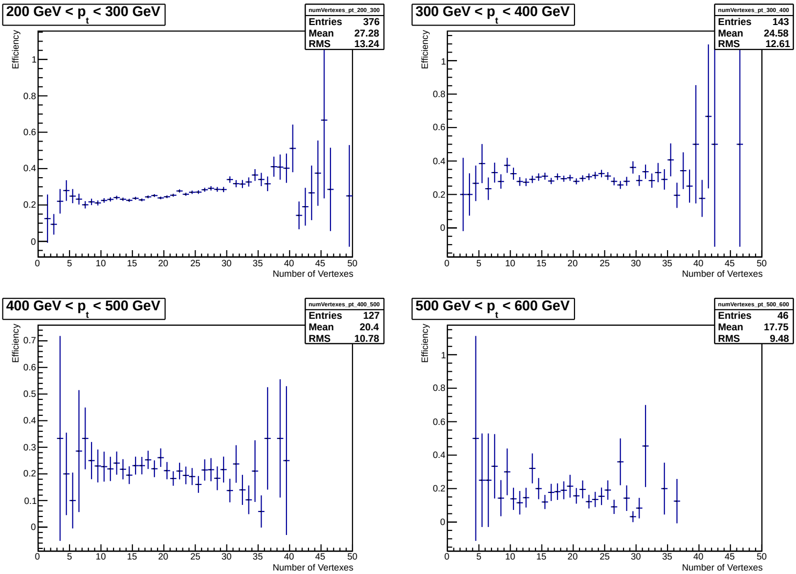




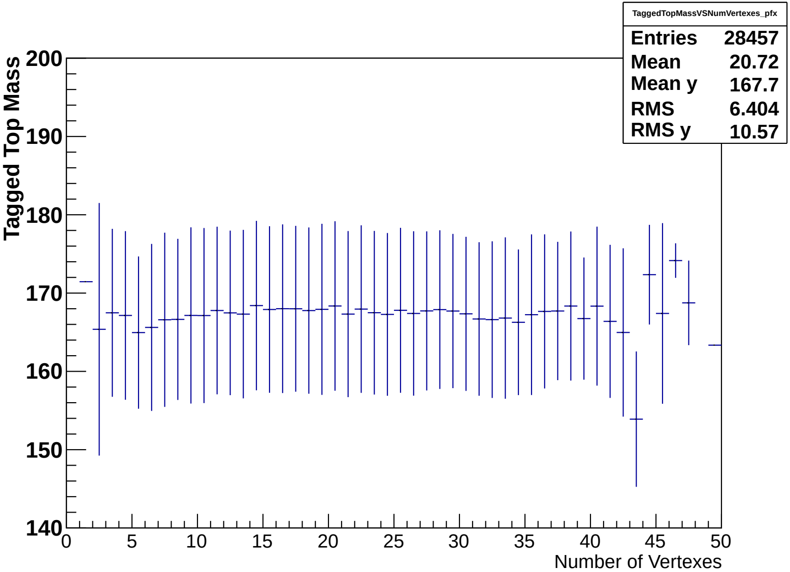




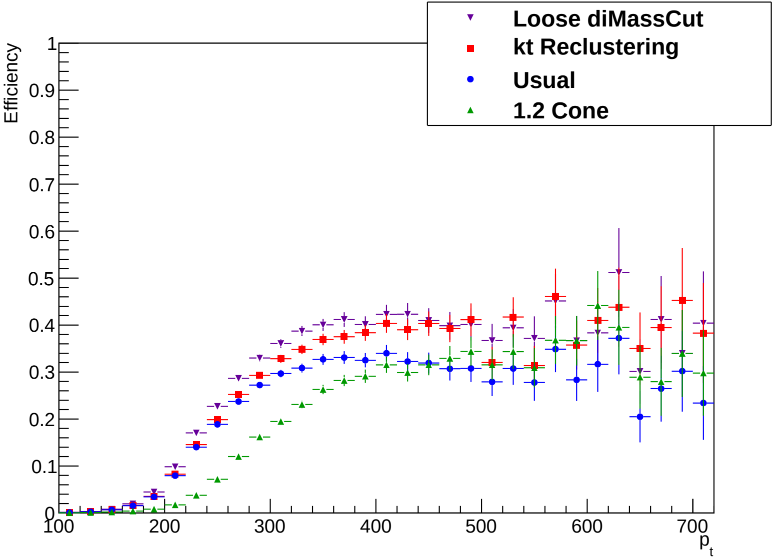
Efficiency pile-up dependence



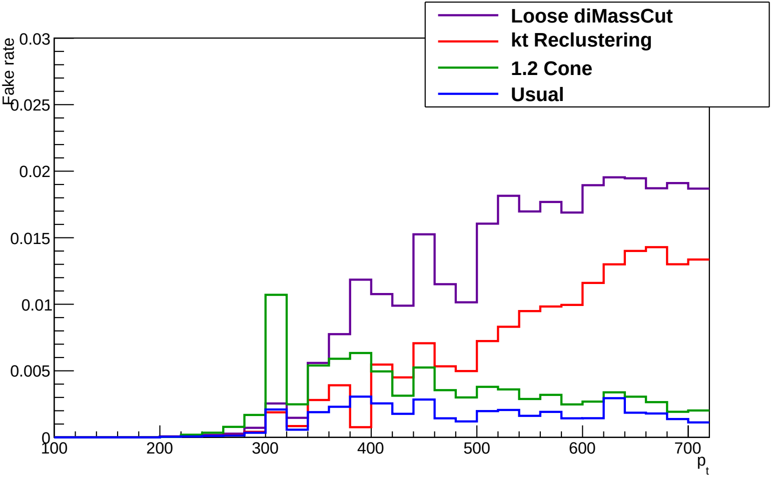
Top mass pile-up dependence



Tuning the HEP Top Tagger: Efficiency

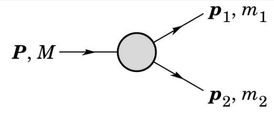


Tuning the HEP Top Tagger: Fake rate



m_T and m_{T2}: definitions

m_T

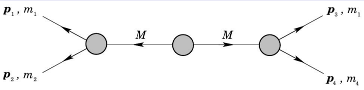


- Used for semi-invisible final state: m_1 visible, m_2 invisible (eg. a neutrino)
- Its distribution can be used to evaluate M. (Used to measure W mass)

$$m_T^2(p_{1T}, m_1, p_{2T}, m_2) = m_1^2 + m_2^2 + 2(E_{1T}E_{2T} - p_{1T}p_{2T})$$

$$\text{where } E_{1T} = \sqrt{m_1^2 + p_{1T}^2}$$

m_{T_2}



Used when a pair of particle decays semi-invisibly: particles 1 and 3 are invisible (LSP) while particles 2 and 4 are visible ($t\bar{t}$).

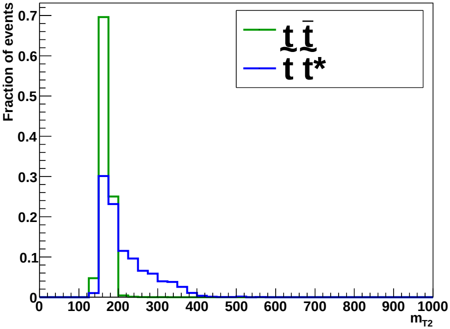
$$m_{T_2}^2(p_{2T}, m_2, p_{4T}, m_4, \not{p}_T, m_1) = \min_{\not{q}_T^{(1)} + \not{q}_T^{(2)} = \not{p}_T} \left[\max\left\{m_T^2(p_{2T}, m_2, \not{p}_T, m_1), m_T^2(p_{4T}, m_4, \not{p}_T, m_1)\right\}\right]$$

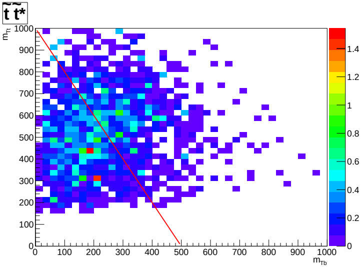
m_T and m_{T_2}: signal selection

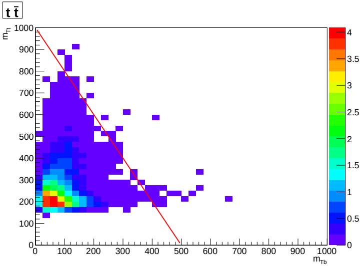
- $m_{Tt} = m_T(p_t, m_t, \not{E}_T, 0)$
- $m_{Tb} = m_T(p_b, m_b, \not{E}_T, 0)$

Cuts

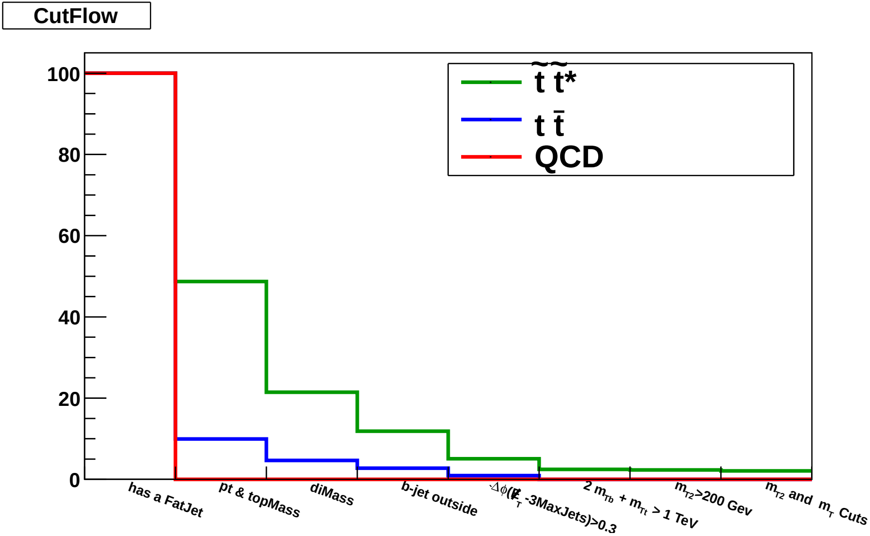
- $m_{T_2} > 200$ GeV
- $2m_{Tb} + m_{Tt} > 1$ TeV







Analysis: a first glimpse

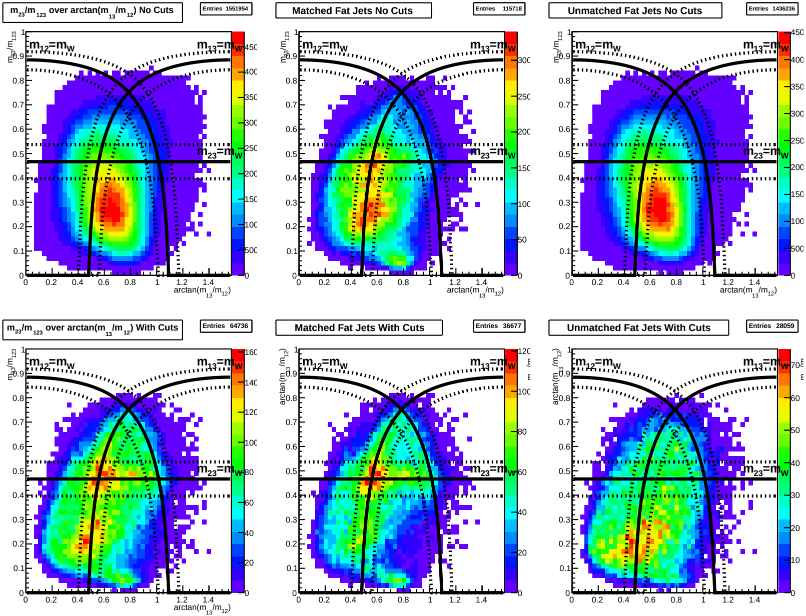


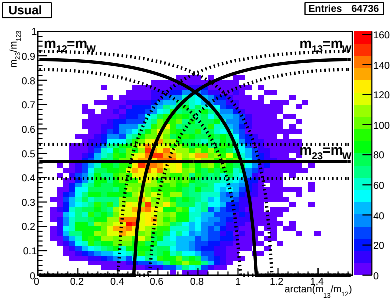
Conclusions

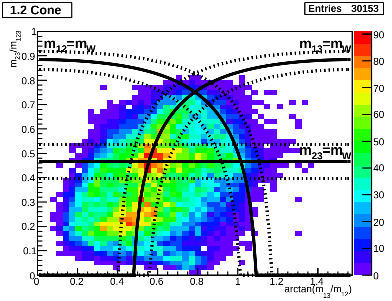
- The Hep Top Tagger is now implemented in the CMS software.
- The tagger's Efficiency and Fake rate were evaluated from Monte Carlo:

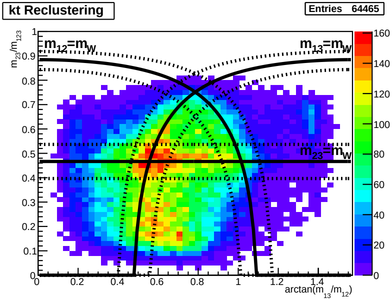
  efficiency: between 20% and 40%
  fake rate: on the order of 0.5%
- Pile-up has little effect on the tagger's performance.
- Looked into several possible improvements.
- We have made modest progression towards a new baseline search for stops.

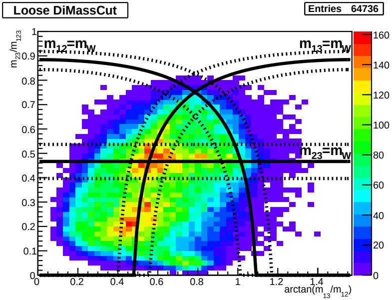
BACKUP

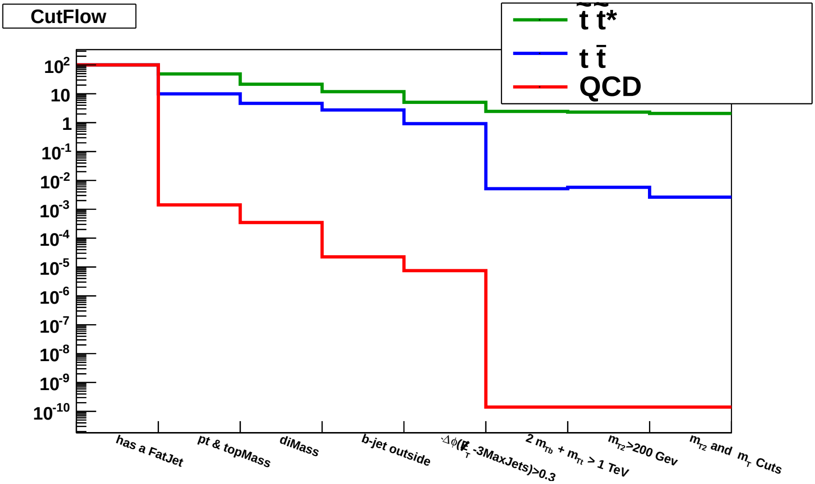

m23/m123 over arctan(m13/m12) No Cuts
Entries 1551954
Matched Fat Jets No Cuts
Entries 115718
Unmatched Fat Jets No Cuts
Entries 1436236
m23/m123 over arctan(m13/m12) With Cuts
Entries 64736
Matched Fat Jets With Cuts
Entries 36677
Unmatched Fat Jets With Cuts
Entries 28059
m12=mW
m13=mW
m23=mW
m23/m123
arctan(m13/m12)


Usual
Entries 64736
$m_{12}=m_W$
$m_{13}=m_W$
$m_{23}=m_W$
m_{23}/m_{123}
arctan(m_{13}/m_{12})
0 0.2 0.4 0.6 0.8 1 1.2 1.4
0 0.1 0.2 0.3 0.4 0.5 0.6 0.7 0.8 0.9 1
0 20 40 60 80 100 120 140 160


1.2 Cone
Entries 30153
$m_{12}=m_W$
$m_{13}=m_W$
$m_{23}=m_W$
m_{23}/m_{123}
arctan(m_{13}/m_{12})
0 0.2 0.4 0.6 0.8 1 1.2 1.4
0 0.1 0.2 0.3 0.4 0.5 0.6 0.7 0.8 0.9 1
0 10 20 30 40 50 60 70 80 90


kt Reclustering
Entries 64465
$m_{12}=m_W$
$m_{13}=m_W$
$m_{23}=m_W$
m_{23}/m_{123}
arctan(m_{13}/m_{12})
0 0.2 0.4 0.6 0.8 1 1.2 1.4
0 0.1 0.2 0.3 0.4 0.5 0.6 0.7 0.8 0.9 1
0 20 40 60 80 100 120 140 160


Loose DiMassCut
Entries 64736
$m_{12}=m_W$
$m_{13}=m_W$
$m_{23}=m_W$
m_{23}/m_{123}
arctan(m_{13}/m_{12})
0 0.2 0.4 0.6 0.8 1 1.2 1.4
0 0.1 0.2 0.3 0.4 0.5 0.6 0.7 0.8 0.9 1
0 20 40 60 80 100 120 140 160


CutFlow
$\tilde{t}\,\tilde{t}^*$
$t\,\bar{t}$
QCD
10^2
10
1
10^{-1}
10^{-2}
10^{-3}
10^{-4}
10^{-5}
10^{-6}
10^{-7}
10^{-8}
10^{-9}
10^{-10}
has a FatJet
pt & topMass
diMass
b-jet outside
$\Delta\phi(\not{E}_T$ -3MaxJets)>0.3
2 m_{T_b} + m_{T_t} > 1 TeV
m_{T2}>200 Gev
m_{T2} and m_T Cuts

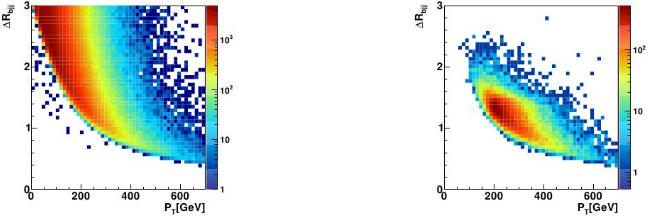
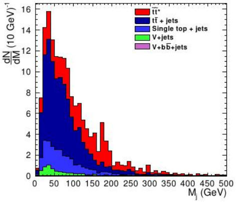


Figure 2: Left: partonic ΔR_{bjj} vs p_T distribution for a Standard Model $t\bar{t}$ sample. Right: the same correlation, but only for tagged top quarks and based on the reconstructed kinematic properties.

Signal: $t\bar{t} + \not{E}_T$

Background

- $t\bar{t}$
  - Where a t decays in b + a hadronic τ + a hard neutrino.
  - Plus $\not{E}_T$ due to poor jet reconstruction .
  - Where a t decays leptonically producing a hard neutrino and the lepton veto fail.
- Z+jets where the Z decays invisibly.
- QCD.



$(m_{\tilde{t}}, m_{\chi_1^0})$= (340 GeV, 0 GeV) at $\sqrt{s}$ =7 TeV and $\mathcal{L} = 5$ fb^{-1}

Signal selection

- Lepton veto.
- $\not{E}_T > 175$ GeV.
- A top-tagged fat-jet.
- Outside the fat jet a b-tagged jet.

HEP Top Tagger

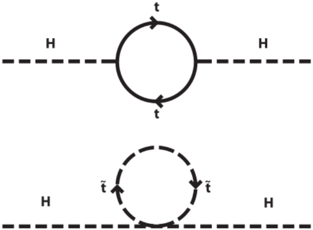
- Start with a Cambridge/Aachen Fat-Jet with $\Delta R = 1.5$ and undo step by step the clustering algorithm to look for subjets.
- When the mother jet becomes two daughters keep both if the larger mass is less than 80% of the mother mass. Otherwise keep the daughter with larger mass.
  Continue as long subjets' mass > 30 GeV. (To cut soft QCD)
- Filter subjets with $\Delta R_{filter} = \min(0.3, \Delta R_{ij}/2)$.
- Keep the five filtered subjets with highest p_T.
- Select the set of three subjets whose combined mass is closest to m_{top}.
- Sort the three selected subjets by p_T (j_1, j_2, j_3) and calculate combined invariant masses m_{12}, m_{13}, m_{23}.
- Check selection requirements.

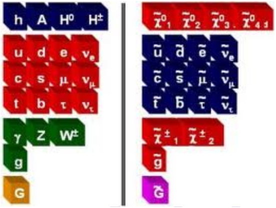
SUSY

Hierarchy problem

The SM Higgs is a fundamental scalar $\Longrightarrow$ his mass is quadratically divergent (with momentum cut off)

- Fine tuning (not "natural")
- Physics BSM that cut off quadratically divergent loop contribution to to Higgs mass
  - SUSY
  - Little Higgs
  - ...





Stop

Top quark is the fermion with the highest Yukawa coupling ⟹ His loop is the one that contribute the most to the Higgs' mass correction.

- In MSSM the running masses of squarks are degenerate at high scale
- Due to large Yukawa coupling stop mass run down at lower energies

⟹ Stop should be the lightest squark

Stop's decay channels

$pp \to \tilde{t}\tilde{t}^*$

Supposing R-parity, stops at LHC should be pair produced.

$\tilde{t} \to t\chi_1^0$

We suppose stop to be sufficiently heavy to decay into a top and a neutralino but not enough to kinematically allows his decay to a top and a gluino.

Other analyses, to avoid QCD background, ask the top to decay semileptonically:

$$pp \to \tilde{t}\tilde{t}^* \to (t\chi_1^0)(\bar{t}\chi_1^0) \to (bl^+\nu\chi_1^0)(\bar{b}jj\chi_1^0) + (\bar{b}l^-\bar{\nu}\chi_1^0)(bjj\chi_1^0)$$

We, complementary, look for a fully hadronic decay:

$$pp \to \tilde{t}\tilde{t}^* \to (t\chi_1^0)(\bar{t}\chi_1^0) \to (bjj\chi_1^0)(\bar{b}jj\chi_1^0)$$