Search for BSM Decaying to Top Quarks

Petar Maksimovic, Johns Hopkins

Jet 1: Top Tagging
pt 589.1 GeV/c,
3 subjets,
mass = 186.7 GeV/c²,
minMass = 87.2 GeV/c²

Jet 2: Jet Pruning
pt 484.3 GeV/c,
mass = 68.8 GeV/c²
Jet 2 + 3: Mass = 167

Jet 3: pt 47.8 GeV/c,
b-tag discriminant 4.2
Why $t\bar{t}$?

- In many BSM models, third generation is special (esp. top)
- Easiest to study the spectrum of $t\bar{t}$ invariant masses, $m_{t\bar{t}}$
  1) Narrow resonance ($\Gamma \ll \text{resolution}$)
  2) Not-so-narrow resonance (use $\Gamma \sim 10\%$ of mass)
  3) Enhancement of the spectrum (no bump)

→ model-independent search for BSM by studying $m_{t\bar{t}}$ distribution!
Three kinematic regimes

Low mass regime:
- isotropic event topology
- `standard' top selection
- combinatorial event reco. (kinematic fit)
- b-tagging

High mass regime:
- relativistic ("boosted") top quarks
- hadronic top decay products (jets) merge
- need to break them into sub-jets ...
- … using dedicated jet clustering tools
- no b-tagging

"Intermediate mass" regime:
- partially merged hadronic top decays
- neither high nor low mass work well
- b-tagging works
“Low mass” lepton+jets

- Event selection identical to $t\bar{t}$ cross-section measurement
  - isolated leptons (both $\mu$ and e), and $E_T^{\text{miss}} > 20$ GeV/$c$
- Top quarks not very energetic $\rightarrow$ no jet merging
  - Events reconstructed using a full kinematic fit
- Backgrounds:
  - QCD from data
    - normalization from template fit to $E_T^{\text{miss}}$
  - All other backgrounds ($t\bar{t}$, W+heavy flavor, W+light flavor) from MC
    - however, MC templates can morph according to systematics
      $\rightarrow$ to a large degree, data driven as well
### "Low mass": systematics

<table>
<thead>
<tr>
<th>Uncertainty</th>
<th>Variation</th>
<th>Type</th>
</tr>
</thead>
<tbody>
<tr>
<td>Luminosity</td>
<td>4.5%</td>
<td>rate</td>
</tr>
<tr>
<td>Electron efficiency (trigger + ID + isolation)</td>
<td>3%</td>
<td>rate</td>
</tr>
<tr>
<td>Muon efficiency (trigger + ID + isolation)</td>
<td>3%</td>
<td>rate</td>
</tr>
<tr>
<td>$\bar{t}t$ cross section</td>
<td>15%</td>
<td>rate</td>
</tr>
<tr>
<td>Single top cross section</td>
<td>30%</td>
<td>rate</td>
</tr>
<tr>
<td>$W/Z+$jets yield</td>
<td>50%</td>
<td>rate</td>
</tr>
<tr>
<td>Drell-Yan yield</td>
<td>30%</td>
<td>rate</td>
</tr>
<tr>
<td>$W/Z+c+X$</td>
<td>100%</td>
<td>rate</td>
</tr>
<tr>
<td>$W/Z+b+X$</td>
<td>100%</td>
<td>rate</td>
</tr>
<tr>
<td>Muon multijet yield</td>
<td>50–75%</td>
<td>rate</td>
</tr>
<tr>
<td>Electron multijet yield</td>
<td>45–70%</td>
<td>rate</td>
</tr>
</tbody>
</table>

Jet energy scale
Jet energy resolution
b tagging efficiency (b jets $p_T < 670$ GeV)
b tagging efficiency (c jets $p_T < 670$ GeV)
b tagging efficiency ($p_T > 670$ GeV)
Mistagging rate
Factorisation scale for $W$ events
$\bar{t}t$ modelling
Factorisation scale for $\bar{t}t$ events
Matching scale for $\bar{t}t$ events
Multiple collisions

$p_T, \eta$ dependent
6–20% per jet
1.6–8% per jet
Twice that for b jets
Twice that at 670 GeV
11%

$\pm 1\sigma$ generator parameters
$\pm 1\sigma$ generator differences
$\pm 1\sigma$ generator parameters
8% inelastic cross section

MC templates can morph in the Likelihood governed by these parameters!
'Low mass' e+jets: data and backgrounds

(μ+jets similar, in backup)

TOP-11-009
"Low mass" lepton+jets: limits

CMS Preliminary
4.7 fb⁻¹ at \(\sqrt{s} = 7\) TeV

Upper Limit \(\sigma_{Z'} \times BR(Z' \rightarrow \bar{t}t)\) [pb]

- CLs Observed (95% CL)
- CLs Expected (95% CL)
- ± 1σ Expected
- ± 2σ Expected
- Topcolor Z' (\(\Gamma/m_{Z'}=1.2\%\))

Upper Limit \(\sigma_{Z'} \times BR(Z' \rightarrow \bar{t}t)\) [pb]

- CLs Observed (95% CL)
- CLs Expected (95% CL)
- ± 1σ Expected
- ± 2σ Expected
- Topcolor Z' (\(\Gamma/m_{Z'}=10\%\))
“Low mass” KK gluon limit

CMS Preliminary
4.7 fb\(^{-1}\) at \(\sqrt{s} = 7\) TeV

Upper Limit \(\sigma_{g'} \times BR(g' \rightarrow t\bar{t})\) [pb] vs. \(m_{KK \text{ gluon}} [\text{TeV/c}^2]\)

- Red dots: CLs Observed (95% CL)
- Black line: CLs Expected (95% CL)
- Light grey shaded region: \(\pm 1\sigma\) Expected
- Dark grey shaded region: \(\pm 2\sigma\) Expected
- Purple line: 1st KK gluon (Agashe et al)
“High mass” lepton+jets

- More boosted, two well-separated hemispheres

- One side: lepton + b-jet, other side: likely merged top

- Trigger:
  - High quality electron, \( p_T^e > 70 \text{ GeV/c} \) and \( |\eta| < 2.5 \)
  - High quality muon, \( p_T^\mu > 35 \text{ GeV/c} \) and \( |\eta| < 2.1 \)

- Basic selection:
  - At least two jets within \( |\eta| < 2.4 \) with \( p_T > 150 \text{ GeV/c} \) and \( p_T > 50 \text{ GeV/c} \)
  - Veto if there is second lepton (e or \( \mu \))
  - “2D cut” (isolation for boosted tops)

\[ H_T^{\text{lep}} \equiv E_T^{\text{miss}} + p_T^\ell > 150 \text{ GeV} \quad (\approx p_T \text{ of } W) \]
Lepton isolation for boosted tops?

- In “low mass” regime, lepton isolation is key to suppress QCD
  - isolation = cut on energy in a fixed cone around the lepton...

- Problem: as $p_T(t)$ increases, b-jet is closer to lepton → inefficient

- Solution: “2D cut”:
  \[
  \Delta R(\ell, \text{closest jet}) > 0.5 \\
  \text{OR} \quad p_T^{\text{rel}}(\ell, \text{closest jet}) > 25 \text{ GeV/c}
  \]
Reconstruction of $m_{t\bar{t}}$

- Reconstruct n's four-vector:
  - set $p_{T,\nu} = $ MET
  - solve quadratic eqn.

\[
p_{z,\nu}^\pm = \frac{\mu p_{z,\ell}}{p_{T,\ell}^2} \pm \sqrt{\frac{\mu^2 p_{z,\ell}^2}{p_{T,\ell}^4} - \frac{E_{\ell}^2 p_{T,\ell}^2}{p_{T,\ell}^2} - \mu^2}
\]

- Select which jet goes with $W$ (lepton + "$\nu$”): minimize

\[
\Delta R_{\text{sum}} = \Delta R(b_1, t_1) + \Delta R(\nu, t_1) + \Delta R(l, t_1)
\]
Event reconstruction in “high mass” analysis

- Top decay products are well separated
- Simple jet-to-parton association is performed
- Works well, but better at higher Z' masses (more back-to-back)
Background determination

- Sample with low $H_T^{lep}$ is signal depleted
  - Template fit of $H_T^{lep}$ determines background normalizations
  - Used to define likelihood in the signal region (high $H_T^{lep}$)

- We fit both simultaneously!
  - Likelihood = product of two likelihoods for 1D distributions of $H_T^{lep}$ and $m_{t\bar{t}}$
  - Template morphing for shape-changing systematics
Result of background “fit”

- In fact, not a fit, but Bayesian marginalization
  - Maxima of posteriors $\iff$ “fit results”
- Dominated by $W$+jets (no b-tagging or top tagging)
Limit in $\mu+$jets

- Only 1.14 fb-1 used!
  - a powerful analysis
- Being updated now, ready by summer
- Bayesian limit
  - CLs for summer

CMS preliminary, $\sqrt{s} = 7$ TeV

Limit on $\sigma(pp \to Z' \to Zt\bar{t})$ [pb]

- Median expected limit
- Observed limit (95% C.L.)
- Central $1\sigma$ expected limit
- Central $2\sigma$ expected limit

- Topcolor $Z'$, 3.0% width, Harris et al.
- Topcolor $Z'$, 1.2% width, Harris et al.
e+jets: data and backgrounds

- Backgrounds normalized from MC
  - W+jets
  - Z+jets
- Constrain norm. within large syst. errors (30%)
- 1D template fit
  (with template morphing)
Limit in e+jets

CMS 2011 Preliminary, $\sqrt{s} = 7$ TeV

$\sigma(Z' \to t\bar{t}) \cdot BR$ [pb]

- expected (95% C.L.)
- central 1$\sigma$ expected
- central 2$\sigma$ expected
- observed (95% C.L.)

Topcolor Z', 1.2% width, Harris et al.
Topcolor Z', 3.0% width, Harris et al.

$L = 4.33$ fb$^{-1}$, e+jets

Z' mass [TeV/c$^2$]
Boosted hadronic top

- Energetic top → jets merge
- Type 1 (single jet):
  - all three jets from top merged into a single jet.
  - JHU TopTagger.
- Type 2 (two jets):
  - two jets from W are contained in a single jet + b-jet separately
  - W tagging plus b tagging.
Top Tagging

- “JHU Top Tagger” (Kaplan, Rehermann, Schwartz, and Tweedie, arXiv:0806.0848), tweaked by CMS

- Cluster jets with Cambridge-Aachen with R=0.8; retrace two steps of clustering sequence back to find subjets

- Three variables
  - jet mass ($\sim m_t$)
  - number of subjets
  - min pairwise mass ($\sim m_W$)

140 < $m_{\text{jet}}$ < 250 GeV/c²

$N_{\text{subjets}} \geq 3$

$m_{\text{min}} > 50$ GeV/c²

CMS Simulation, $\sqrt{s} = 7$ TeV

[Graph showing probability distribution of m(top jet)]
W tagging

- Jet pruning
  - Ellis, Vermillion, Walsh (arXiv:0903.5081)
  - Improves mass resolution by removing soft, large angle particles

- Undo last step of jet clustering (also CA 0.8) to find two subjets. Identify W's with:
  - “Mass drop”,
  - Pruned jet mass

\[
\mu \equiv \frac{m_{\text{leading sub jet}}}{m_{\text{jet}}} < 0.4
\]

\[60 < m_{\text{jet}} < 130 \text{GeV/c}^2\]
Validating jet substructure in data

- In “low mass” lepton+jets sample, look for merged W's
- Test MC (Madgraph + Pythia Z2) → work surprisingly well!
Validating jet substructure in data (2)

- Use W peak from W-tagged jets for 
  "substructure energy scale" = 1.01 ± 0.01
- Also measure efficiency correction for MC = 0.97 ± 0.03
$Z' \rightarrow t\bar{t}$ in all-hadronic: two event topologies

- Aimed at higher $Z'$ masses
- Type 1+1
  - Jet $p_T > 350$ GeV/c
  - Events with two top tags
- Type 1+2
  - Jet $p_T > 350, 200, 30$ GeV/c
  - $140 < m_{\text{jet}} < 250$ GeV/c$^2$
  - Top tag in one hemisphere
  - $W$ tag + $b$-tag on the other
Example of a “Type 1+2” candidate

Type 1 top: Top-Tagged jet

Type 2 top:
- W-tag
- b-tag

Jet 1: Top Tagging
pt 589.1 GeV/c,
3 subjets,
mass = 186.7 GeV/c2,
minMass = 87.2 GeV/c2

Jet 2: Jet Pruning
pt 484.3 GeV/c,
mass = 68.8 GeV/c2
Jet 2 + 3: Mass = 167

Jet 3: pt 47.8 GeV/c,
b-tag discriminant 4.2
Mistag probability for Top Tagging

- Probability to mis-identify a non-top QCD jet as a top tag
- Obtained from “substructure sideband” in Type 1+2 events
- Apply to pre-top-tag $m_{t\bar{t}}$ distribution to get non-top background
All-hadronic: data + background prediction

CMS Preliminary

L = 4.6 fb^{-1}, \sqrt{s} = 7 TeV  

Type 1+1

- Observed
- Non-top multijet
- $t\bar{t}$ simulation
- $Z'(1 \text{ TeV/c}^2) \sigma = 1.0 \text{ pb}$
- $Z'(1.5 \text{ TeV/c}^2) \sigma = 0.18 \text{ pb}$
- $Z'(2 \text{ TeV/c}^2) \sigma = 0.06 \text{ pb}$
- $Z'(3 \text{ TeV/c}^2) \sigma = 0.03 \text{ pb}$

Events / 100 GeV/c^2

Data – Predicted

\tilde{t} mass (GeV/c^2)

CMS Preliminary

L = 4.6 fb^{-1}, \sqrt{s} = 7 TeV  

Type 1+2

- Observed
- Non-top multijet
- $t\bar{t}$ simulation
- $Z'(1 \text{ TeV/c}^2) \sigma = 1.0 \text{ pb}$
- $Z'(1.5 \text{ TeV/c}^2) \sigma = 0.18 \text{ pb}$
- $Z'(2 \text{ TeV/c}^2) \sigma = 0.06 \text{ pb}$
All-hadronic: limit for narrow width

CMS Preliminary

EXO-11-006
All-hadronic: limit for 10% width

L = 4.6 fb\(^{-1}\) at \(\sqrt{s} = 7\) TeV

CMS Preliminary

EXO-11-006
All-hadronic: limit for KK Gluon

\[ L = 4.6 \text{ fb}^{-1} \text{ at } \sqrt{s} = 7 \text{ TeV} \]

- **Observed (95% CL)**
- **Expected (95% CL)**
  - \( \pm 1 \text{ s.d. Expected} \)
  - \( \pm 2 \text{ s.d. Expected} \)

KK Gluon, Agashe et al
All-hadronic: general enhancement

- What if there is no peak, just change in $m_{t\bar{t}}$ spectrum?

$$ S = \frac{\int_{1\text{TeV}}^{+\infty} \left( \frac{d\sigma_{SM+NP}}{dm_{t\bar{t}}} \right) dm_{t\bar{t}}}{\int_{1\text{TeV}}^{+\infty} \left( \frac{d\sigma_{SM}}{dm_{t\bar{t}}} \right) dm_{t\bar{t}}} $$

- Set limit on

- Assume same efficiency as SM $t\bar{t}$ production

- Integrate $m_{t\bar{t}} > 1 \text{TeV}/c^2$ correct for smearing in $m_{t\bar{t}}$

<table>
<thead>
<tr>
<th></th>
<th>1+1</th>
<th>1+2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Expected SM $t\bar{t}$ events</td>
<td>$194 \pm 106$</td>
<td>$129 \pm 80$</td>
</tr>
<tr>
<td>Expected non-top multijet events</td>
<td>$1546 \pm 45$</td>
<td>$2271 \pm 130$</td>
</tr>
<tr>
<td>Total expected events</td>
<td>$1740 \pm 115$</td>
<td>$2400 \pm 153$</td>
</tr>
<tr>
<td>Observed events</td>
<td>$1738$</td>
<td>$2423$</td>
</tr>
<tr>
<td>$t\bar{t}$ efficiency</td>
<td>$(2.5 \pm 1.3) \times 10^{-4}$</td>
<td>$(1.6 \pm 1.0) \times 10^{-4}$</td>
</tr>
</tbody>
</table>

- Counting experiment gives CLs limit: $S < 2.6$ @ 95% C.L.
Summary

- CMS searches for $t\bar{t}$ resonances are in full swing
- Already reached sub-picobarn limits (depending on the model)
  - individual searches rule out $Z'/KKg$ below 1.2-2 TeV
- For summer – easy:
  - update $\mu$+jets to 4.6 fb-1
  - combine high mass measurements (ready to do so)
- For summer – may or may not be easy:
  - add top tagging to lepton+jets analyses
  - add b-tagging to both lepton+jets and all-hadronic
- In any case, 2012 will be fun!
B A C K U P
"Low mass" $\mu$+jets

CMS Preliminary
4.7 fb$^{-1}$ at $\sqrt{s} = 7$ TeV
$\mu$, = 3 jets, $\geq$ 1 b-tags

CMS Preliminary
4.7 fb$^{-1}$ at $\sqrt{s} = 7$ TeV
$\mu$, $\geq$ 4 jets = 0 b-tags

CMS Preliminary
4.7 fb$^{-1}$ at $\sqrt{s} = 7$ TeV
$\mu$, $\geq$ 4 jets, = 1 b-tag

CMS Preliminary
4.7 fb$^{-1}$ at $\sqrt{s} = 7$ TeV
$\mu$, $\geq$ 4 jets, $\geq$ 2 b-tags
e+jets: further cuts to suppress QCD

- $E_T^{\text{miss}} > 50$ GeV

and the “triangular cuts”:

- $\Delta\phi(e, E_T^{\text{miss}}) < \frac{1.5}{75} E_T^{\text{miss}} + 1.5$,
  $\Delta\phi(e, E_T^{\text{miss}}) > -\frac{1.5}{75} E_T^{\text{miss}} + 1.5$,

- $\Delta\phi(j_1, E_T^{\text{miss}}) < \frac{1.5}{75} E_T^{\text{miss}} + 1.5$,
  $\Delta\phi(j_1, E_T^{\text{miss}}) > -\frac{1.5}{75} E_T^{\text{miss}} + 1.5$
e+jets: QCD suppressed!

- QCD estimated from a 1D template fit of MET
- QCD template = invert electron ID and triangular cuts

\[ \text{\textbullet} \quad \text{QCD is small} \]

\[ \Rightarrow \text{After all the cuts, QCD is small} \]
### “High mass” systematics (e+jets)

<table>
<thead>
<tr>
<th>Source of systematic uncertainty</th>
<th>Uncertainty</th>
<th>Type</th>
</tr>
</thead>
<tbody>
<tr>
<td>tt cross section</td>
<td>15%</td>
<td>Rate</td>
</tr>
<tr>
<td>Single top cross section</td>
<td>15%</td>
<td>Rate</td>
</tr>
<tr>
<td>W+jets cross section</td>
<td>30%</td>
<td>Rate</td>
</tr>
<tr>
<td>Z+jets cross section</td>
<td>30%</td>
<td>Rate</td>
</tr>
<tr>
<td>QCD multijet</td>
<td>50%</td>
<td>Rate</td>
</tr>
<tr>
<td>Luminosity</td>
<td>4.5%</td>
<td>Rate</td>
</tr>
<tr>
<td>Trigger</td>
<td>4%</td>
<td>Rate</td>
</tr>
<tr>
<td>Jet Energy Scale</td>
<td>±1\sigma(p_T, \eta)</td>
<td>Rate &amp; Shape</td>
</tr>
<tr>
<td>Scale ((Q^2 = M(t)^2 + \sum p_T(jet)^2))</td>
<td>(2Q^2) and (0.5Q^2)</td>
<td>Rate &amp; Shape</td>
</tr>
<tr>
<td>Matching</td>
<td>2 and (0.5 \times ) default</td>
<td>Rate &amp; Shape</td>
</tr>
<tr>
<td>Pileup</td>
<td>±1\sigma</td>
<td>Rate &amp; Shape</td>
</tr>
</tbody>
</table>