Application of MVA in new physics searches
Lesya Shchutska (UF)
MLHEP School
St. Petersburg, Russia
CMS in the cavern before closing
- $4\pi$ general purpose detector: registers all particles in transverse plane
- spectrometer with a large magnetic field (3.8 T): precise measurement of charged particles
- electromagnetic calorimeter for photon/electron measurements
- hadron calorimeter with moderate energy resolution: $\sim 10\%$ above 100 GeV
- muon system: identification (also improves momentum resolution for TeV muons)
Success of the Standard model

Top Physics

Single top

EWK Physics

1. $\sigma_{t\bar{t}} = 241.5 \pm 1.4\text{(stat.)} \pm 5.7\text{(syst.)} \pm 6.2\text{(lumi.)} \text{ pb}$

2. $m_t = 172.38 \pm 0.14\text{(stat.)} \pm 0.64\text{(syst.)} \text{ GeV}$

Stunning agreement with SM expectations!
Moving from discovery to precision measurement!
One important very basic symmetry (have not been seen (yet?))

- for each $\frac{1}{2}$-integer spin particle (fermion) there is an integer spin partner (boson) and vice versa
  - complete spectrum of partners to standard model particles
  - their spins are different by $\frac{1}{2}$ unit
  - they are heavier (or else we’d have seen them already).

Supersymmetry introduces partners to each SM particle with $\Delta(\text{spin}) = \frac{1}{2}$.

Wide variety of signatures and rich phenomenology.

Higgs bosons.
SUSY implications

SUSY unifies the strengths of all forces at $\sim 10^{16}$ GeV
SUSY implications

Explains 25% of the energy in the universe: the **dark matter**
Beautiful but not minimal theory

- SUSY is a broken symmetry: masses of superpartners are not fixed by theory
- a parameter space which is impossible to fully exclude but to only constrain

Within the MSSM only:

- MSSM: 109 parameters
- pMSSM: 19 parameters
- CMSSM: 5 parameters

Complementary strategies are required to maximally constrain the parameters:

- direct and indirect dark matter detection experiments
- study of the rates of the rare processes (e.g. heavy-flavor physics)
- precision SM production cross section measurements (e.g. $t\bar{t}$ production)

- direct SUSY particle production in the pp collisions at the LHC and their detection in ATLAS or CMS experiments
Challenges

- SUSY particle production cross sections are 10 and more orders of magnitude lower than full pp collision rate at the LHC.

- comparable to other rare SM processes, e.g.:
  - \( \sigma(t\bar{t}W) = 203 \text{ fb} \)
  - \( \sigma(t\bar{t}Z) = 206 \text{ fb} \)
  - \( \sigma(t\bar{t}H) = 129 \text{ fb} \)
  - expect around 4000 events produced in full 8 TeV dataset

- compared to \( \sigma(t\bar{t}) = 252 \text{ pb} \): 3 orders of magnitude to go

- \( t\bar{t}+X \) processes are great proxies and test-cases for possible SUSY discovery
  - \( t\bar{t}W \) and \( t\bar{t}Z \) story is told in the next slides
Discovering new processes in standard model

The same full 8 TeV dataset (∼20 fb⁻¹), and the search for \( tt\bar{Z} \) and \( ttW \).

In CMS:
1. June 30, 2014: “Measurement of top quark-antiquark pair production in association with a W or Z boson in pp collisions at \( \sqrt{s} = 8 \) TeV”
2. May 08, 2015: “Measurement of top quark pair production in association with a W or Z boson using event reconstruction techniques”

And in ATLAS:
1. July 4, 2014: “Evidence for the associated production of a vector boson (W, Z) and top quark pair in the dilepton and trilepton channels in pp collision data at \( \sqrt{s} = 8 \) TeV collected by the ATLAS detector at the LHC”
2. July 22, 2015: “Measurement of the \( ttW \) and \( ttZ \) production cross sections in pp collisions at \( \sqrt{s} = 8 \) TeV with the ATLAS detector”

What is the difference between the two?
First analyses by CMS and ATLAS at 7 and 8 TeV used a cut-based approach in the most sensitive channels, observed \( t\bar{t}Z \) at \( \sim 3\sigma \) significance and achieved \( \sim 2\sigma \) sensitivity to \( t\bar{t}W \)

<table>
<thead>
<tr>
<th>Channels used</th>
<th>Process</th>
<th>Cross section</th>
<th>Significance</th>
</tr>
</thead>
<tbody>
<tr>
<td>2( \ell )</td>
<td>( t\bar{t}W )</td>
<td>( 170^{+50}<em>{-40}(\text{stat.})^{+40}</em>{-30}(\text{syst.}) ) fb</td>
<td>1.6 ( \sigma )</td>
</tr>
<tr>
<td>3( \ell + 4\bar{\ell} )</td>
<td>( t\bar{t}Z )</td>
<td>( 200^{+50}<em>{-40}(\text{stat.})^{+40}</em>{-30}(\text{syst.}) ) fb</td>
<td>3.1 ( \sigma )</td>
</tr>
</tbody>
</table>

First results

\( \sqrt{s} = 8 \text{ TeV}, L_{\text{int}} = 19.5 \text{ fb}^{-1} \)

\( \int L dt = 20.3 \text{ fb}^{-1} \)

**CMS**

**ATLAS**

\( \sqrt{s} = 8 \text{ TeV} \)

\( \sigma / \sigma_{\text{SM}}(t\bar{t}W) \)

\( \sigma / \sigma_{\text{SM}}(t\bar{t}Z) \)
ttZ signatures

- explore a presence of a Z boson:
  - use all final states with \( Z \rightarrow \ell\ell (\ell = e \text{ or } \mu) \)
  - and a fact of having b-jets, missing transverse energy (\( E_T^{\text{miss}} \)), and/or light flavor jets
- can use 2\( \ell \)OS, 3\( \ell \), 4\( \ell \) signatures
  - first CMS C&C analysis used only 3\( \ell \), 4\( \ell \)
  - first ATLAS paper covered 2\( \ell \)OS, 2\( \ell \)SS and 3\( \ell \)
ttW signatures

- explore leptonic W decays
- CMS looked at $2\ell SS$
- ATLAS used $2\ell SS$ and $3\ell$ final states
- in addition: b jets, light flavor jets, $E_T^{\text{miss}}$

- Best S/B in same-sign $2\ell$ (SS) and $3\ell$ final states
- Expect leptons, b-jets, and missing energy, plus an extra lepton or extra light flavor jets

Diagram:

- $\bar{q}$
- $W^+$
- $e^+, \mu^+$
- $b$
- $SS$
- $3l$
SM processes with fewer particles (and orders of magnitude larger cross sections):
- $t\bar{t}$, $Z$, $WZ$, $ZZ$

to fake $t\bar{t}V$ need to have extra objects:
- either leptons - from b decays or misidentified jets
- or jets - from initial state radiation (ISR) or pile up

these backgrounds have different importance in various search channels:
- first categorize in lepton multiplicity
How it looks like: same-sign dileptons

**Preselection**

- lepton kinematics: leading lepton $p_T$ is harder in signal
- hadronic activity: more jets in signal $\implies$ larger $H_T = \sum_{\text{jets}} p_T^{\text{jet}}$
- $t$ quark mass reconstruction of 3 jets: poor mass resolution

**Full selection**
Same-sign dileptons: result

End up with a selection like:

- 2 same-sign leptons with \( p_T > 40 \) GeV
- at least 3 jets with \( p_T > 30 \) GeV, at least one of them a b jet
- hadronic activity \( H_T > 155 \) GeV

\[ \Rightarrow \text{box-like cuts} \]

- expect around 15 signal events in all categories with about 25\( \pm \)7 background
- while initially have \( 4000 \times 0.25 \times 0.25 = 250 \) 2\( \ell \)SS signal \( t\bar{t}W \) events
- need to cut very hard to arrive at reasonable S/B ratio

<table>
<thead>
<tr>
<th></th>
<th>( \mu^+\mu^+ )</th>
<th>( e^+\mu^+ )</th>
<th>( e^+e^+ )</th>
<th>( \mu^-\mu^- )</th>
<th>( e^-\mu^- )</th>
<th>( e^-e^- )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( t\bar{t}W )(expected)</td>
<td>2.8 ± 0.4</td>
<td>5.1 ± 0.5</td>
<td>2.2 ± 0.3</td>
<td>1.1 ± 0.2</td>
<td>2.3 ± 0.3</td>
<td>1.0 ± 0.2</td>
</tr>
<tr>
<td>Misidentified lepton</td>
<td>1.0 ± 0.6</td>
<td>4.1 ± 2.1</td>
<td>1.6 ± 0.9</td>
<td>0.7 ± 0.4</td>
<td>3.0 ± 1.5</td>
<td>1.7 ± 0.9</td>
</tr>
<tr>
<td>Mismeasured charge</td>
<td>—</td>
<td>0.4 ± 0.1</td>
<td>0.7 ± 0.2</td>
<td>—</td>
<td>0.4 ± 0.1</td>
<td>0.7 ± 0.2</td>
</tr>
<tr>
<td>Irreducible</td>
<td>0.7 ± 0.4</td>
<td>1.6 ± 0.9</td>
<td>0.9 ± 0.5</td>
<td>0.5 ± 0.3</td>
<td>1.4 ± 0.7</td>
<td>0.7 ± 0.4</td>
</tr>
<tr>
<td>WZ</td>
<td>0.1 ± 0.1</td>
<td>0.4 ± 0.1</td>
<td>0.1 ± 0.1</td>
<td>0.1 ± 0.1</td>
<td>0.4 ± 0.1</td>
<td>0.2 ± 0.1</td>
</tr>
<tr>
<td>( t\bar{t}Z )</td>
<td>0.6 ± 0.3</td>
<td>0.9 ± 0.5</td>
<td>0.5 ± 0.3</td>
<td>0.4 ± 0.2</td>
<td>1.0 ± 0.5</td>
<td>0.5 ± 0.3</td>
</tr>
<tr>
<td>Total background</td>
<td>2.4 ± 0.7</td>
<td>7.4 ± 2.3</td>
<td>3.9 ± 1.1</td>
<td>1.7 ± 0.5</td>
<td>6.1 ± 1.8</td>
<td>3.7 ± 1.1</td>
</tr>
<tr>
<td>Total expected</td>
<td>5.2 ± 0.8</td>
<td>12.5 ± 2.4</td>
<td>6.1 ± 1.1</td>
<td>2.8 ± 0.5</td>
<td>8.4 ± 1.8</td>
<td>4.7 ± 1.1</td>
</tr>
<tr>
<td>Observed</td>
<td>6</td>
<td>12</td>
<td>5</td>
<td>1</td>
<td>6</td>
<td>6</td>
</tr>
</tbody>
</table>

- similar approach in other lepton multiplicities
Final result

- very much compatible with SM
- evidence for $t\bar{t}Z$ production
- quite low sensitivity to $t\bar{t}W$

<table>
<thead>
<tr>
<th>Channels used</th>
<th>Process</th>
<th>Cross section</th>
<th>Significance</th>
</tr>
</thead>
<tbody>
<tr>
<td>$2\ell SS$</td>
<td>$t\bar{t}W$</td>
<td>$170^{+90}_{-80} \pm 70 \text{ fb}$</td>
<td>1.6</td>
</tr>
<tr>
<td>$3\ell + 4\ell$</td>
<td>$t\bar{t}Z$</td>
<td>$200^{+80+40}_{-70-30} \text{ fb}$</td>
<td>3.1</td>
</tr>
<tr>
<td>$2\ell SS + 3\ell + 4\ell$</td>
<td>$t\bar{t}W + t\bar{t}Z$</td>
<td>$380^{+100+80}_{-90-70} \text{ fb}$</td>
<td>3.7</td>
</tr>
</tbody>
</table>
First ATLAS analysis

- use 2\(\ell\)SS, 2\(\ell\)OS, 3\(\ell\) channels
  - 3\(\ell\) and 2\(\ell\)SS: comparable signal and background \(\mapsto\) C\&C
  - opposite-sign dilepton: small signal in huge background \(\mapsto\) neural network
    - a channel was not present in previous C\&C search!

- main sensitivity of the search is from cut\&count analyses
  - combined result is comparable with CMS cut\&count paper

<table>
<thead>
<tr>
<th>Process</th>
<th>Signal Strength</th>
<th>Observed (\sigma)</th>
<th>Expected (\sigma)</th>
</tr>
</thead>
<tbody>
<tr>
<td>(t\bar{t}V)</td>
<td>(0.91^{+0.27}_{-0.24})</td>
<td>4.6</td>
<td>4.6</td>
</tr>
<tr>
<td>(t\bar{t}W)</td>
<td>(1.31^{+0.62}_{-0.50})</td>
<td>3.0</td>
<td>2.3</td>
</tr>
<tr>
<td>(t\bar{t}Z)</td>
<td>(0.72^{+0.33}_{-0.28})</td>
<td>2.8</td>
<td>3.4</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Process</th>
<th>Signal Strength</th>
<th>Observed (\sigma)</th>
<th>Expected (\sigma)</th>
</tr>
</thead>
<tbody>
<tr>
<td>(t\bar{t}V)</td>
<td>(0.77^{+0.63}_{-0.56})</td>
<td>1.4</td>
<td>1.7</td>
</tr>
<tr>
<td>(t\bar{t}W)</td>
<td>(0.57^{+2.48}_{-2.30})</td>
<td>0.3</td>
<td>0.4</td>
</tr>
<tr>
<td>(t\bar{t}Z)</td>
<td>(0.77^{+0.69}_{-0.59})</td>
<td>1.4</td>
<td>1.5</td>
</tr>
</tbody>
</table>

Simultaneous fit of two signal strengths in all channels

<table>
<thead>
<tr>
<th>Channel</th>
<th>(\mu_{t\bar{t}Z})</th>
<th>(\mu_{t\bar{t}W})</th>
<th>Observed (\sigma)</th>
<th>Expected (\sigma)</th>
</tr>
</thead>
<tbody>
<tr>
<td>trilepton and same-sign dilepton</td>
<td>(0.70^{+0.30}_{-0.28})</td>
<td>(1.37^{+0.62}_{-0.51})</td>
<td>4.1</td>
<td>4.1</td>
</tr>
<tr>
<td>opposite-sign dilepton</td>
<td>(0.77 \pm 0.65)</td>
<td>(0.71 \pm 2.41)</td>
<td>0.4</td>
<td>0.6</td>
</tr>
<tr>
<td>combination</td>
<td>(0.71^{+0.28}_{-0.26})</td>
<td>(1.30^{+0.59}_{-0.48})</td>
<td>4.4</td>
<td>4.4</td>
</tr>
</tbody>
</table>
†V reapproached at CMS

Instead of relying on simple kinematic properties of objects in the event:

1. **employ physics knowledge**: be aware of the presence of top quarks
   - reconstruct a t\(\bar{t}\) pair in t\(\bar{t}Z\), t\(\bar{t}W\) and t\(\bar{t}\) events

2. **acknowledge complicated kinematics**: use multivariate techniques
   - construct a linear discriminant (**MatchLD**) with reconstructed masses and object properties
   - perform a training on the t\(\bar{t}\) (background) MC

**Note**: cannot easily separate the gain between the two

---

**Event reconstruction**

- **ttZ (3l)**
  - \(t\bar{t}Z (3l)\)
  - \(QQ\)
  - Mass, charge
  - b-tag, charge
  - \(M_T\)

- **ttbar (SS)**
  - \(t\bar{t}\) (SS)
  - Mass, charge
  - b-tag, charge
  - \(M_T\)
  - Mass
Event reconstruction: full input list to MatchLD

- extensively use the knowledge of event topology
- prefer a lot of non-linear and physics-motivated variables and a linear discriminant
  - some inputs are another MVA results, e.g. b-tagging score
- rather than fewer simple 4-vectors and a deep learning neural network

<table>
<thead>
<tr>
<th>Variable</th>
<th>OS ttZ</th>
<th>SS ttW and 3ℓ ttZ</th>
<th>3ℓ ttW</th>
</tr>
</thead>
<tbody>
<tr>
<td>b-jet CSV</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
</tr>
<tr>
<td>Higher jet CSV from W</td>
<td>Y</td>
<td>Y</td>
<td>N</td>
</tr>
<tr>
<td>Lower jet CSV from W</td>
<td>Y</td>
<td>Y</td>
<td>N</td>
</tr>
<tr>
<td>Leptonic top b-jet charge</td>
<td>N</td>
<td>Y</td>
<td>N</td>
</tr>
<tr>
<td>Hadronic top b-jet charge</td>
<td>N</td>
<td>Y</td>
<td>N</td>
</tr>
<tr>
<td>Top b-jet charge</td>
<td>Y</td>
<td>N</td>
<td>Y</td>
</tr>
<tr>
<td>Anti-top b-jet charge</td>
<td>Y</td>
<td>N</td>
<td>Y</td>
</tr>
<tr>
<td>Sum of charges of jets from W</td>
<td>Y</td>
<td>Y</td>
<td>N</td>
</tr>
<tr>
<td>Mass of lepton and b-jet from leptonic top</td>
<td>N</td>
<td>Y</td>
<td>N</td>
</tr>
<tr>
<td>Mass of lepton and b-jet from top</td>
<td>N</td>
<td>N</td>
<td>Y</td>
</tr>
<tr>
<td>Mass of lepton and b-jet from anti-top</td>
<td>N</td>
<td>N</td>
<td>Y</td>
</tr>
<tr>
<td>$M_T$ of $E_T^{miss}$ and lepton and b-jet from leptonic top</td>
<td>N</td>
<td>Y</td>
<td>N</td>
</tr>
<tr>
<td>Mass of two jets from W</td>
<td>Y</td>
<td>Y</td>
<td>N</td>
</tr>
<tr>
<td>Mass of b-jet and one jet from W from hadronic top</td>
<td>Y</td>
<td>Y</td>
<td>N</td>
</tr>
<tr>
<td>Mass of b-jet and two jets from W from hadronic top</td>
<td>Y</td>
<td>Y</td>
<td>N</td>
</tr>
<tr>
<td>Ratio of $M_T$ to mass for jets from top or W</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
</tr>
</tbody>
</table>
Example in $t\bar{t}Z$

Example of constructing a MatchLD score:

- left: $W \rightarrow q\bar{q}$ dijet mass distribution for correctly picked jets in $t\bar{t}$
- center: dijet mass distribution for any pair of jets in the $t\bar{t}$ event
- right: scaled ratio of the two - used as a score

- Renormalize ratio histogram to have average value of 1.0 for events where the jets are correctly assigned
- Attempt every permutation of jets matched to the decay products of the $t\bar{t}$ system. Choose permutation with highest score $= \log(\prod_i \text{ratio}_i)$

Application in data:

- test every permutation of jets and leptons with different “mother” hypotheses
- for each permutation take the product of values from ratio histograms: $\log(\prod_i \text{ratio}_i)$
  - this score has median at 1 for correct match, and $< 1$ for wrong reconstruction
- the highest MatchLD score out of all permutation is used as another discriminating variable

---

**Figure:**

Histograms showing dijet mass distributions for matched and unmatched jets, along with ratio histograms for $W$ mass.

- $t\bar{t} \rightarrow b\bar{v}lqq$ for matched jets
- $t\bar{t} \rightarrow b\bar{v}lq$ for unmatched jets
- $W$ mass distribution

- Ratio histograms for good and bad matches.
for events where both b’s from the top and both q’s from the W are reconstructed as jets, 75% of 4 jet and 40% of ≥ 5 jet events have every jet correctly matched to its parent particle

• for correct matches, the average ratio is 1 (so the match score centers near 0)

• partial matches (all but one jet matched) allow to identify signal events where one of the quarks forms a jet outside our acceptance
MatchLD itself is passed over to the BDT:

- train boosted decision tree with ttZ vs. WZ and ttbar MC
- also include $M_T$ of system, and partial matches to ttZ system

- $ttZ \rightarrow \ell^+\ell^- (bl\nu)(bq)$
- $ttZ \rightarrow \ell^+\ell^- (l\nu)(bqq)$
- $ttZ \rightarrow \ell^+\ell^- (bl\nu)(qq)$
• train boosted decision trees with \( \ttW \) vs. \( \tt \) MC

• use a mix of kinematic and matching variables
  • kinematic variables from \( p_T \) of objects and jet b-tag (CSV)
  • event matching variables for \( \ttW \) and \( \tt \) systems

• reconstructing events with a linear discriminant first allows for more input variables and better separation than a BDT alone, since BDTs are more limited by the statistics of training events
BDT output

**ttW**

- **Fakes**
- **Q flip**
- **WZ**
- **Other**
- **ttH**
- **ttZ**
- **ttW**

**SS $\mu\mu$, $\geq 4$ jets**

**SS $e\mu$, $\geq 4$ jets**

**SS $ee$, $\geq 4$ jets**

**ttZ**

- **3l, $\geq 4$ jets**
- **4l, $\geq 1$ jet**

*Final discriminants*

$SS \& ee, \geq 4$ jets

$SS \& $$, \geq 4$ jets

$SS \& e$, $\geq 4$ jets

Andrew Brinkerhoff
## Results

### ttZ:

<table>
<thead>
<tr>
<th>Channels</th>
<th>Expected</th>
<th>Observed</th>
<th>Expected</th>
<th>Observed</th>
<th>Significance</th>
</tr>
</thead>
<tbody>
<tr>
<td>OS</td>
<td>$206^{+142}_{-118}$</td>
<td>$257^{+158}_{-129}$</td>
<td>$1.0^{+0.72}_{-0.57}$</td>
<td>$1.25^{+0.76(+1.76)}_{-0.62(-1.16)}$</td>
<td>1.84</td>
</tr>
<tr>
<td>3ℓ</td>
<td>$206^{+79}_{-63}$</td>
<td>$257^{+85}_{-67}$</td>
<td>$1.0^{+0.42}_{-0.32}$</td>
<td>$1.25^{+0.45(+1.02)}_{-0.36(-0.62)}$</td>
<td>4.55</td>
</tr>
<tr>
<td>4ℓ</td>
<td>$206^{+153}_{-109}$</td>
<td>$228^{+150}_{-107}$</td>
<td>$1.0^{+0.77}_{-0.53}$</td>
<td>$1.11^{+0.76(+1.79)}_{-0.52(-0.86)}$</td>
<td>2.65</td>
</tr>
<tr>
<td>All</td>
<td>$206^{+62}_{-52}$</td>
<td>$242^{+65}_{-55}$</td>
<td>$1.0^{+0.34}_{-0.27}$</td>
<td>$1.18^{+0.35(+0.79)}_{-0.29(-0.51)}$</td>
<td>5.73</td>
</tr>
</tbody>
</table>

### ttW:

<table>
<thead>
<tr>
<th>Channels</th>
<th>Expected</th>
<th>Observed</th>
<th>Expected</th>
<th>Observed</th>
<th>Significance</th>
</tr>
</thead>
<tbody>
<tr>
<td>SS</td>
<td>$203^{+88}_{-73}$</td>
<td>$414^{+135}_{-112}$</td>
<td>$1.0^{+0.45}_{-0.36}$</td>
<td>$2.04^{+0.74(+1.52)}_{-0.61(-1.05)}$</td>
<td>3.44</td>
</tr>
<tr>
<td>3ℓ</td>
<td>$203^{+215}_{-94}$</td>
<td>$210^{+225}_{-203}$</td>
<td>$1.0^{+1.09}_{-0.96}$</td>
<td>$1.03^{+1.07(+2.39)}_{-0.99(-1.92)}$</td>
<td>1.03</td>
</tr>
<tr>
<td>All</td>
<td>$203^{+84}_{-71}$</td>
<td>$382^{+117}_{-102}$</td>
<td>$1.0^{+0.43}_{-0.35}$</td>
<td>$1.88^{+0.66(+1.35)}_{-0.56(-0.95)}$</td>
<td>3.54</td>
</tr>
</tbody>
</table>

More specific and targeted approach makes a difference between the *evidence* and *discovery*!

1. **ttZ**: from 3.1 (3.1)$\sigma$ to 5.7(6.4)$\sigma$
2. **ttW**: from 2.0 (1.6)$\sigma$ to 3.5(4.8)$\sigma$
What about SUSY?
Again top-quark: superpartner

\( \tilde{t} \) cancels out the largest divergency in the Higgs boson mass - from \( t \) quark:

1. 1-loop order: top contribution corrected by \( \tilde{t} \rightarrow m_{\tilde{t}} \approx \mathcal{O}(100 \text{ GeV}) \)

\[
\Delta m_{H}^{2} \propto - \frac{1}{H} \quad + \quad \frac{1}{H}
\]

2. 2-loop order: gluino enters \( \tilde{t} \) mass \( \rightarrow m_{\tilde{g}} \approx \mathcal{O}(1 \text{ TeV}) \)

3. not too heavy \( \tilde{b}_{L} \): in the doublet with \( \tilde{t}_{L} \)

Plethora of ATLAS and CMS analyses looking for gluinos, 3rd generation squarks and gauginos!
Topologies for Top squark search

Landscape is thoroughly combed in search of an elusive top-quark partner - needed to stabilize a Higgs boson mass:

Decays in 2-particle spectrum: $\tilde{t}_1$ and $\tilde{\chi}_1^0$

$$\Delta m(\tilde{t}_1, \tilde{\chi}_1^0) < 0$$

Also look at 3-particle spectra:

$\tilde{t}_1, \tilde{\chi}_1^0, \tilde{\chi}_1^0$: $\tilde{t}_1 \rightarrow b\tilde{\chi}_1^\pm (\rightarrow W\tilde{\chi}_1^0)$

$\tilde{t}_1, \tilde{\chi}_2^0, \tilde{\chi}_1^0$: $\tilde{t}_1 \rightarrow t\tilde{\chi}_2 (\rightarrow Z/H\tilde{\chi}_1^0)$

$\tilde{t}_2, \tilde{t}_1, \tilde{\chi}_1^0$: $\tilde{t}_2 \rightarrow Z/H\tilde{t}_1 (\rightarrow t\tilde{\chi}_1^0)$

Each final state is targeted by one or more dedicated heavily optimized searches

Individual analyses are combined to estimate sensitivity in mixed scenarios
Top squark searches in two plots

- probed phase space extends up to 750 GeV in $\tilde{t}$ mass and 260 GeV in $\tilde{\chi}^0_1$ mass
- invent new tools to access difficult regions: $m_{\tilde{t}} - m_{\tilde{\chi}^0_1} \approx m_t$ or $m_W$

![Diagram showing top squark production and sensitivities](image)

- exploit cascade decays of heavier $\tilde{t}_2$ or $\tilde{g}$ to $\tilde{t}_1$
- design monojet (+soft leptons) to catch $\tilde{t} \rightarrow c \tilde{\chi}^0_1$ ($\tilde{t} \rightarrow b f f'$)

![Diagram showing comprehensive combination of results](image)

- sensitivity from $t\bar{t}$ spin correlation

---


32/50
Search strategy

Multiple signal regions target different signal kinematics based on $m(\tilde{t})$ and $m(\tilde{\chi}_1^0)$

**'low $\Delta m'$**
Best sensitivity with looser requirements → small systematic uncertainty on the background is key

**'high $\Delta m'$**
Best sensitivity with tighter requirements background suppressed to O(10) events → low statistics regime
Performed analysis in two different ways in terms of how the signal regions are defined

- **Cut-based**: “Square” cuts on several variables
- **BDT**: Combine variables into a BDT and define signal region by cutting on the output

Multiple signal regions to target different signal kinematics
Example variables

- $M_{T2}$
- $M_W$
- $W$, $T$, $M$
- 100, 150, 200, 250, 300, 350, 400, 450, 500 GeV
- $L_{dt} = 19.5 \text{ fb}^{-1}$
- $s = 8 \text{ TeV}$
- CMS Simulation
- Preselection + $\mathcal{E}_T > 120 \text{ GeV}$

**Example Diagrams**

1. **$m_{top}$**
   - $M_W^{W}$
   - $M_{T2}$
   - JHEP07 (2012) 110

2. **Leading b-jet $p_T$**
   - Signal $\times 100$
   - Background $\times 100$
   - Fraction of $H_T$ in same hemisphere as MET

3. **$t$-channel diagrams**
   - $t\to Wb$
   - $t\to \nu b$
   - $W$-boson
   - $\nu$

4. **$E_T^{miss}$**
   - $\tilde{t}$
   - $\tilde{\chi}_1^0$
   - $W$
   - $b$-jet

5. **$p_T(b_1)$ [GeV]**
   - Symbol representation of $p_T$
Example BDT output discriminant used to define the signal regions
## Signal regions definitions

### Signal variables and regions used in two analyses

<table>
<thead>
<tr>
<th>Selection</th>
<th>BDT</th>
<th>cut-based</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Low $\Delta M$</td>
</tr>
<tr>
<td>$E_T^{\text{miss}}$ (GeV)</td>
<td>yes</td>
<td>$&gt; 150, 200$, $250, 300$</td>
</tr>
<tr>
<td>$M_T^W$ (GeV)</td>
<td>yes</td>
<td></td>
</tr>
<tr>
<td>$\text{min} \Delta \phi$</td>
<td>yes</td>
<td>$&gt; 0.8$</td>
</tr>
<tr>
<td>$H_T^{\text{ratio}}$</td>
<td>yes</td>
<td></td>
</tr>
<tr>
<td>hadronic top $\chi^2$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>leading b-jet $p_T$ (GeV)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\Delta R(\ell, \text{leading b-jet})$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>lepton $p_T$</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note: for cut-based there is a low $\Delta m$ and high $\Delta m$ selection
Signal Region Choice for Interpretation

Map of best signal region on m(\(\tilde{t}\)) vs m(\(\tilde{\chi}_1^0\)) plane → signal regions selected using best expected limit

(These correspond roughly to the regions defined to target different signals)
See a result of a clean comparison: MVA-based search goes to the corners not accessible to C&C:

low signal efficiency/high level of background ones!
Cut-based vs. BDT Result Comparison

Ratio of expected cross section upper limit: BDT / cut-based

BDT is $\sim 10\text{-}30\%$ better than cut-and-count in most of parameter space
Another $\tilde{t}$ example

Targeted $\tilde{t}$ search pushes the boundaries:

- Aims at all-hadronic final state for $\tilde{t} \to t\tilde{\chi}_1^0$ and $\tilde{t} \to bW\tilde{\chi}_1^0$
- Operates in a $\tilde{t}$-search region with a range of boosts:
  - From unboosted top quarks at low $\Delta M$
  - To merged top quarks at very high $\Delta M$
- For this developed a new top quark reconstruction algorithm with varying cone size

A set of MVA is trained for different kinematic regimes
- Achieved sensitivity surpasses the one from previous $0\ell+1\ell$ combined: in both $\tilde{t}$ and $\tilde{\chi}_1^0$ masses
- Up to 775 GeV for $\tilde{t}$ mass with $m_{\tilde{\chi}_1^0} < 200$ GeV
- Up to 275 GeV for $\tilde{\chi}_1^0$ mass
More general new physics search?

How do we define what we are looking for?

- in general, the tails of distributions:
  - high object multiplicity, high $E_T^{\text{miss}}$, $H_T$ etc
- stealth SUSY:
  - rely on uncovering tags: hard initial state radiation, vector boson fusion topologies
- non-standard detector signatures:
  - multiply charged particles
  - disappearing tracks
  - displaced tracks
- anything else unthought of?

General question:

- how do we quantify what we see when applying complicated techniques
We ask the network: “Whatever you see there, I want more of it!”

This creates a feedback loop: if a cloud looks a little bit like a bird, the network will make it look more like a bird. This in turn will make the network recognize the bird even more strongly on the next pass and so forth, until a highly detailed bird appears, seemingly out of nowhere.
Plenty of “new discoveries” in line!

if apply advanced tools mindlessly
Resorting to lower levels

- detector response
- object reconstruction
- particle identification
- kinematic selection

- the higher in the pyramid the more process-dependent a task is
- if target general new physics cannot be too specific in kinematics
- but if new physics decays to SM particles can use all advanced techniques there
- such hidden MVA are all over the place: electrons, muons, b-jet identification
- ensures low level of irreducible background
in practice we perform a sort of pattern recognition:

- combine information from different subdetectors
- and form a picture about which particles flew through a detector

a separate optimization is done for each particle type based on physics knowledge
Example: electrons and photons

- Cluster reconstruction in ECAL
  - Common for both electrons and photons (electrons also reconstructed as photons)
  - Designed to collect bremsstrahlung and conversions in extended phi region

- Dedicated track reconstruction for electrons
  - “Gaussian sum” filter allows for tracks with large bremsstrahlung

- A range of detector-based variables to separate real electrons from misID jets
  - Put together either in cut-based ID or MVA (BDT) ID: MVA can have 10-40% better efficiency with the same background rejection
  - Huge gain in efficiency especially in multilepton events
Taus - even more complicated objects

- not a stable particles: decays within the detector acceptance
- reconstructed in individual decay modes: higher level object
- requires both charged hadrons and electromagnetic objects
- series of MVA are trained to distinguish $\tau$ from jets or leptons:
  - automatically propagated to all analyses using $\tau$

$\tau \rightarrow \pi \nu$  
$\tau \rightarrow a_1 \nu$  
$\tau \rightarrow \rho \nu$

1-prong  
3-prong  
1-prong+strip

decays within the detector acceptance
reconstructed in individual decay modes: higher level object
requires both charged hadrons and electromagnetic objects
series of MVA are trained to distinguish $\tau$ from jets or leptons:
  - automatically propagated to all analyses using $\tau$

$\tau \rightarrow \pi \nu$  
$\tau \rightarrow a_1 \nu$  
$\tau \rightarrow \rho \nu$

1-prong  
3-prong  
1-prong+strip

configured to $\tau$ decays

Data  
Simulation  
$\text{TauES}^{*1.03}$  
$\text{TauES}^{*0.97}$

visible $\tau$ mass (GeV/c$^2$)
Conclusions

- To achieve a discovery big reduction rate of the backgrounds are needed
- Important to use all the features of our data to discriminate signal from background
- MVA are cautiously used in the new physics searches:
  - and usually only in well-motivated and understood scenarios
- But at the same time employed at many levels within the HEP framework:
  - Event level: Higgs searches, top events
  - Cone level: Tau vs jet reconstruction
  - Track level: particle identification
  - Lifetime and flavour tagging: b-tagging
- For a discovery, new tools are a two-way street:
  - can be used to enhance a cut-based hint
  - if signal seen first in MVA search: a cut-based analysis later to confirm the result
Plenty of discoveries are ahead: hopefully, mostly real ones!