Systematically Searching for New Physics at the LHC

Daniel Whiteson
UC Irvine

UC Davis HEP Seminar, April 2014
Outline

I. Dark Matter
II. Topological Models
III. Deep networks
What do we know?

unknown unknown

known

unknown

known unknown

known known
Colliders: true alchemy

We can create new forms of matter, even if we have little or no idea of what we are looking for!
Simultaneously, we know we know very little...

So we know where to look.

And, we built this new collider.

The magic of a collider is that you can make kinds of matter that you don't have around. It's this amazing quantum mechanical magic that's just turned on.

These two things are coming together right now.
Important assumption: Requires some interaction with SM.
DM @ Colliders
Look everywhere

Mono-jet most powerful for $qqXX$

Each mode has unique models where it is a discovery mode.
Outline

A. Mono-W
B. Mono-Z
C. Mono-Higgs
Mono-W
Mono-W theory

Searches with Mono-Leptons

Yang Bai\textsuperscript{a,b} and Tim M.P. Tait\textsuperscript{c}

1208.4361
Mono-jet

Missing Momentum

$q/g$
Mono-heavy jet

1309.4017 (PRL)

fat jet

sub-jet

sub-jet

W/Z

Missing Momentum

Ning Zhou, UCI
mono-W, etc

Fat jet $p_T > 250$ 1309.4017 (PRL)
two subjets giving $m_{\text{jet}} = [50, 120]$
No e, mu, gamma
$\leqslant 1$ additional narrow jets
MET $> 350$ or 500
Limits

1309.4017 (PRL)
FIG. 4: Limits on $\Lambda$ as a function of $m_\chi$. 18
"Indirect" is an excluded region which is a combination of exclusions from the LAT line search, the LAT dwarf bounds and (at higher m_ch) the VERITAS Segue bounds. It is assumed that this DM makes up 100% of cosmological DM, no matter what its annihilation cross section is.
EFTs

(a) Feynman diagram showing an ISR operator.

(b) Feynman diagram showing a $ZZ\chi\chi$ operator.

1404.0051

Andy Nelson, UCI
Simplified models

1404.0051
### Selection

Two OS SF leps mll in [76,106] veto jets, 3rd lep MET angle cuts $|p_T^{\ell\ell} - E_T^{\text{miss}}|/p_T^{\ell\ell} < 0.5$

<table>
<thead>
<tr>
<th>Process</th>
<th>$E_T^{\text{miss}}$ threshold [GeV]</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>150</td>
</tr>
<tr>
<td>$ZZ$</td>
<td>41 ± 15</td>
</tr>
<tr>
<td>$WZ$</td>
<td>8.0 ± 3.1</td>
</tr>
<tr>
<td>$WW, t\bar{t}, Z \rightarrow \tau^+\tau^-$</td>
<td>1.9 ± 1.4</td>
</tr>
<tr>
<td>$Z$+jets</td>
<td>0.1 ± 0.1</td>
</tr>
<tr>
<td>$W$+jets</td>
<td>0.5 ± 0.3</td>
</tr>
<tr>
<td>Total</td>
<td>52 ± 18</td>
</tr>
<tr>
<td>Data</td>
<td>45</td>
</tr>
</tbody>
</table>
Data
Limits....

\[ \int L = 20.3 \text{ fb}^{-1} \quad \sqrt{s} = 8 \text{ TeV} \]

ATLAS

\[ M_\ast \text{ [GeV]} \]

\[ 10^2 \quad 10^3 \quad 10^4 \]

\[ 0 \quad 200 \quad 400 \quad 600 \quad 800 \quad 1000 \]

\[ m_\chi \text{ [GeV]} \]

1404.0051

D1
D5
D9
ZZ\chi\chi \text{ max. } \gamma
ZZ\chi\chi \text{ no } \gamma
Limits....
Limits....
Mono-Higgs

1312.2592
FIG. 1: Schematic diagram for mono-Higgs production in $pp$ collisions mediated by electroweak bosons ($h, Z, \gamma$) or new mediator particles such as a $Z'$ or scalar singlet $S$. The gray circle denotes an effective interaction between DM, the Higgs boson, and other states.
Models: EFT

\[ \lambda |H|^2 |\chi|^2 \]

\[ \frac{1}{\Lambda} |H|^2 \bar{\chi}\chi, \quad \frac{1}{\Lambda} |H|^2 \bar{\chi}i\gamma_5\chi \]

Scalar wimp

Fermion wimp
Vertices

**di-Higgs**

(1) $h \rightarrow XX$ limited by invisible Higgs for $m_X < m_h/2$

(2) For large coupling, $h \rightarrow XX$ grows, suppresses SM $H$ decays!

**4-point vertex**

**Off-shell s-channel Higgs**
Allow ZhXX-like vertices

\[ \frac{1}{\Lambda^2} \chi^\dagger i\partial^\mu \chi H^\dagger iD_\mu H \]  
Scalar wimp

\[ \frac{1}{\Lambda^4} \bar{\chi} \gamma^\mu \chi B_{\mu\nu} H^\dagger D^\nu H \]  
Fermion wimp
Simplified models: vector

(a) \[ q \xrightarrow{Z'} h \xrightarrow{Z'} \chi \]

(b) \[ q \xrightarrow{Z} h \xrightarrow{Z'} \chi \]

with and without Z-Z' mixing
Simplified models: scalar

Box implemented as effective vertex in madgraph
\( m_x = 1 \text{ GeV} \)

\( m_x = 1 \text{ TeV} \)

EFTs  Simp. models
**Gamma-gamma**

**Selection**
- two photons
- \( m_{\gamma\gamma} \) in \([110-130]\)
- MET > 100, 250 (8,14 TeV)

**Backgrounds**
- \( h\rightarrow\gamma\gamma + \text{fake MET} \)
- \( \gamma\gamma + \text{fake MET} \)
- \( Z\gamma\gamma, Z\rightarrow\nu\nu \)
- \( Zh, Z\rightarrow\nu\nu + Wh, W\rightarrow\ell\nu \)
Assuming $h \rightarrow SM$ rates are unchanged
Assuming $h \rightarrow SM$ rates are unchanged
**Note:**
for $m_x < m_h/2$, no valid limits.
Large Lambda boosts $h \rightarrow XX$, suppresses $h \rightarrow$ visible
FIG. 22: Projected LHC mono-Higgs sensitivities at $\sqrt{s} = 8$ TeV (20 fb$^{-1}$) and 14 TeV (300 fb$^{-1}$), with $\gamma\gamma + E_T$ final states, on simplified models. All constraint contours exclude larger couplings or mixing angles. Shaded region is excluded based on perturbativity arguments or requiring $\sin \theta \leq 1$; orange contour denotes limit from invisible $h$ decays; purple contours are exclusion limits from LUX.
**ATLAS**

- **7 TeV γ+MET** (1209.4625)
- **W→jj +MET** (1309.4017)
- **Invisible Higgs** (1402.3244)
- **Z+MET** (1404.0051)
- **W→lν +MET** (soon)

**Pheno**

- **monoZ** (1212.3352)
- **DM combo** (1302.3619)
- **Fermi/LHC** (1307.5064)
- **DM future** (1307.5327)
- **H+MET** (1312.2592)
- **Indirect WW** (1403.6734)
- **Compressed spectra** (forthcoming)
- **mono-Z’** (forthcoming)
Outline

I. Dark Matter
II. Topological Models
III. Deep networks
Searching for new physics

Specific  →  Model  ←  General

Search strategy
Traditional approach

**Bet on a specific theory**
Optimize analysis to squeeze out maximal sensitivity to new physics.

(param 3-N fixed at arbitrary choices)
Model independent search

Discard the model
compare data to standard model

“Never listen to theorists. Just go look for it.”
–A. Pierce, 2010
Compromise

Admit the need for a model
New signal requires a coherent physical explanation, even trivial or effective

Generalize your model
Construct simple models that describe classes of new physics which can be discovered at the LHC.

What are we good at discovering?
Admit the need for a model
New signal requires a coherent physical explanation, even trivial or effective

Generalize your model
Construct simple models that describe classes of new physics which can be discovered at the LHC.

What are we good at discovering? Resonances!
Is this being done?

\[ \ell^+ \]

\[ \ell^- \]

\[ j \]

\[ j \]
Is this being done?

\[ W' \rightarrow Z \rightarrow \ell^+ \ell^- \]

\[ m_{\|} = m_Z \]

\[ m_{jj} = m_W \]
What about this?

$W'$

$\chi_1 \rightarrow \ell^+ \ell^-$

$\chi_2 \rightarrow j j$

$m_{\ell\ell} \neq m_Z$

$m_{jj} \neq m_W$
Missed resonances?

Easy-to-find resonances may exist in our data and nobody has looked!
Missed resonances?

Easy-to-find resonances may exist in our data and nobody has looked!
Systematically Searching for New Resonances at the Energy Frontier using Topological Models

Mohammad Abdullah,¹ Eric Albin,¹ Anthony DiFranzo,¹ Meghan Frate,¹ Craig Pitcher,¹ Chase Shimmin,¹ Suneet Upadhyay,¹ James Walker,¹ Pierce Weatherly,¹ Patrick J. Fox,² and Daniel Whiteson¹

¹Department of Physics and Astronomy, University of California, Irvine, CA 92697
²Fermi National Accelerator Laboratory, Batavia, IL 60510

We propose a new strategy to systematically search for new physics processes in particle collisions at the energy frontier. An examination of all possible topologies which give identifiable resonant features in a specific final state leads to a tractable number of ‘topological models’ per final state and gives specific guidance for their discovery. Using one specific final state, $\ell\ell jj$, as an example, we find that the number of possibilities is reasonable and reveals simple, but as-yet-unexplored, topologies which contain significant discovery potential. We propose analysis techniques and estimate the sensitivity for $pp$ collisions with $\sqrt{s} = 14$ TeV and $L = 300$ fb$^{-1}$. 
Topological models

For a given final state (e.g., $lljj$) construct all models with resonances. Then look for them!
Connections to EFT, Simp. Models

- Full Theories
- Effective Operators
- Simplified models

Mass scale vs. Completeness
Connections to EFT, Simp. Models

- Full Theories
- Effective Operators
- Simplified models
- Topo models

Mass scale vs. Completeness
Mono-Z’

\[ m_{ij} = m_W \text{ or } m_Z \quad \text{and} \quad m_{II} = m_Z \]
Mono-Z’

$m_{||} = m_W$ or $m_Z$

$m_{||} = m_Z$

What about other values?
**Signature**

Heavy resonance + MET
Outline

I. Dark Matter
II. Topological Models
III. Deep networks
How to find NP

Isolate some feature in which two theories \( \text{SM}, \text{SM}+\text{X} \) can be best distinguished.

The data can tell us which hypothesis is preferred via a likelihood ratio:

\[
\frac{L_{\text{SM}+\text{X}}}{L_{\text{SM}}} \quad \frac{P(\text{data} \mid \text{SM}+\text{X})}{P(\text{data} \mid \text{SM})}
\]

Standard Model
- \( \text{SM}+\text{X} \)
- Collider Data

some feature
e.g.
But...

Reality is more complicated.

The full space can be very high dimensional.

Calculating likelihood in $d$-dimensional space requires $\sim 100^d$ MC events.
Neural networks can learn these shapes in high-dim and summarize in a 1D output.
Neural Networks

Essentially a functional fit with many parameters

**Function**
Each neuron’s output is a function of the weighted sum of inputs.

**Goal**
find set of weights which give most useful function

**Learning**
give examples, back-propagate error to adjust weights
Neural Networks

Essentially a functional fit with many parameters

Problem:
Networks with > 1 layer are very difficult to train.

Consequence:
Networks are not good at learning non-linear functions. (like invariant masses!)

In short:
Can’t just throw 4-vectors at NN.
Can’t just use 4v

Can’t give it too many inputs

Painstaking search through input feature space.

Table 3: Discriminating variables used for each channel and category. The filled circles identify which variables are used in each decay mode. Note that variables such as $\Delta R(\tau, \tau)$ are defined either between the two leptons, between the lepton and $\tau_{\text{had}}$, or between the two $\tau_{\text{had}}$ candidates, depending on the decay mode.
Can’t just use 4v

Can’t give it too many inputs

Painstaking search through input feature space.

Also true for BDTs, SVNss, etc

<table>
<thead>
<tr>
<th>Variable</th>
<th>VBF</th>
<th>Boosted</th>
</tr>
</thead>
<tbody>
<tr>
<td>$m_{\text{T}}$</td>
<td>🄽</td>
<td>🄽</td>
</tr>
<tr>
<td>$\Delta R(\ell, \ell)$</td>
<td>🄽</td>
<td>🄽</td>
</tr>
<tr>
<td>$n_{\text{T}}$</td>
<td>🄽</td>
<td>🄽</td>
</tr>
<tr>
<td>$p_{\ell_1}^T$</td>
<td>🄽</td>
<td>🄽</td>
</tr>
<tr>
<td>$p_{\ell_2}^T$</td>
<td>🄽</td>
<td>🄽</td>
</tr>
<tr>
<td>min($\Delta R(\ell, \ell)$)</td>
<td>🄽</td>
<td>🄽</td>
</tr>
<tr>
<td>$\ell_1 \times \ell_2$</td>
<td>🄽</td>
<td>🄽</td>
</tr>
<tr>
<td>$\ell$</td>
<td>🄽</td>
<td>🄽</td>
</tr>
<tr>
<td>$\tau_1, 2$</td>
<td>🄽</td>
<td>🄽</td>
</tr>
</tbody>
</table>

Table 3: Discriminating variables used for each channel and category. The filled circles identify which variables are used in each decay mode. Note that variables such as $\Delta R(\tau, \tau)$ are defined either between the two leptons, between the lepton and $\tau_{\text{had}}$, or between the two $\tau_{\text{had}}$ candidates, depending on the decay mode.
New tools let us train deep networks.

How well do they work?
Real world applications

Head turn: DeepFace uses a 3-D model to rotate faces, virtually, so that they face the camera. Image (a) shows the original image, and (g) shows the final, corrected version.
Deep Learning in High-Energy Physics: Improving the Search for Exotic Particles

P. Baldi, P. Sadowski, and D. Whiteson

1Dept. of Computer Science, UC Irvine, Irvine, CA 92617
2Dept. of Physics and Astronomy, UC Irvine, Irvine, CA 92617

arXiv: 1402.4735
In revision at Nature Comm.
Can deep networks automatically discover useful variables?
4-vector inputs

21 Low-level vars
jet+lepton mom. (3x5)
missing ET (2)
jet btags (4)

Not much separation visible in 1D projections
4-vector inputs

7 High-level vars

\[ m(WWbb) \]
\[ m(Wbb) \]
\[ m(bb) \]
\[ m(bjj) \]
\[ m(jj) \]
\[ m(lv) \]
\[ m(blv) \]
4-vector inputs

7 High-level vars

- \( m(WWbb) \)
- \( m(Wbb) \)
- \( m(bb) \)
- \( m(bjj) \)
- \( m(jj) \)
- \( m(lv) \)
- \( m(blv) \)
4-vector inputs

7 High-level vars
m(WWbb)
m(Wbb)
m(bb)
m(bjj)
m(jj)
m(lv)
m(blv)
Standard NNs

Results
Adding hi-level boosts performance
Better: lo+hi-level.

Conclude:
NN can’t find hi-level vars.
Hi-level vars do not have all info
Standard NNs


Conclude: NN can’t find hi-level vars. Hi-level vars do not have all info.

Also true for BDTs, SVNs, etc.
Deep Networks

Results
Lo+hi = lo.

Conclude:
DN can find hi-level vars.

Hi-level vars do not have all info are unnecessary.
Deep Networks

Results
DN > NN

Conclude:
DN does better than human assisted NN
The AIs win
Results

Identified example benchmark where traditional NNs fail to discover all discrimination power.

Adding human insight helps traditional NNs.

DN not as reliant on signal features. Cuts into background space.
Can Deep Networks help us find SUSY in the data?
Low-level variables
High-level variables

Axial-MET
Met-rel
MT2
Razor
Super-razor
SUSY results

DN doesn’t need help

Outperforms human assisted NN

Margin is smaller
-> high level variables are less helpful and less needed!

<table>
<thead>
<tr>
<th>Technique</th>
<th>Low-level</th>
<th>High-level</th>
<th>Complete</th>
</tr>
</thead>
<tbody>
<tr>
<td>NN</td>
<td>6.5σ</td>
<td>6.2σ</td>
<td>6.9σ</td>
</tr>
<tr>
<td>DN</td>
<td>7.5σ</td>
<td>7.3σ</td>
<td>7.6σ</td>
</tr>
</tbody>
</table>
Preliminary: h->tautau
Preliminary: $h \rightarrow \tau \tau \tau \tau$
Preliminary: $h \rightarrow \tau \tau \tau$

The diagram shows a comparison of Discovery Significance [$\sigma$] for shallow and deep networks with different input types.

- **Shallow networks**:
  - Raw inputs: 1.7
  - Human-assisted: 2.0
  - All input: 2.2

- **Deep networks**:
  - Raw inputs: 2.3
  - Human-assisted: 2.5
  - All input: 2.5

The results indicate that deep networks generally provide higher Discovery Significance compared to shallow networks, with the all input category achieving the highest values.
Summary

**Dark matter:**
- broad-based attack on all LHC signals

**Topological models:**
- Strategy to build complete set of models with discoverable resonances

**Deep networks:**
- Networks can take 4-vectors, find powerful discriminants