Hadronically Identified and Generated, after a Generation

Phil Harris (MIT)
After a 25 year hiatus, rekindling the standard model with low mass hadronic decays

Phil Harris (MIT)
1966
Long Island
New York
Di muon measurements

Lederman's dimuon search on 33 GeV AGS

Lederman's Shoulder

Brookhaven, NY
Searching for the Shoulder

- Boston 1972

Sam Ting proposes to revisit the Lederman Shoulder with electrons at a high resolution
What happened?

- November revolution

\[ e^+e^- \rightarrow \psi @ SLAC \]

Nobel prize 1976

\[ pp \rightarrow J @ Brookhaven \]
The Fallout (From Lederman himself)

• We set up this ingeniously stupid detector in order to study muon pairs. [...] the simple Brookhaven apparatus was so dumb that there was no way you could change it to improve the resolution even a little bit.

• “..After the shock of seeing Ting’s data, I sent around a note saying that any apparatus that can convert this towering peak to our mound of rubble should be proscribed by SALT [Nuclear Arms] talks”
5 Years ago @ CERN

We discovered the Higgs
Unfortunately

- Higgs turned up to be more of what we expected

5 years of the Higgs have largely confirmed the predictions for the Higgs

No hints new physics in the Higgs production
What if we haven't looked finely

Something inconsistent?
What if we haven't looked finely

What's a minuteman doing here?
The beginning of a revolution?
New Technology?

• This talk is a story of
  - How can probe the Higgs at high $p_T$ for the first time

• It is also a story of in short of new tech...yielding
  - Adapted jet substructure to large background
  - A new approach to measure resonances
  - The measurement of $pp \rightarrow W \rightarrow qq$ production for 1$^{st}$ time
  - The measurement of $pp \rightarrow Z \rightarrow bb$ in 1 jet for 1$^{st}$ time
The quest for the $W/Z \rightarrow qq$ peak
What's the first goal?

Find:

A W boson decaying to quarks
In an experiment

Find:

A $W$ boson decaying to quarks
A low $p_T$ we resolve them as jets
In an experiment

Find:

A $W$ boson decaying to quarks
A low $p_T$ we resolve them as jets
Was last observed 30 years
Result was not very convincing
Have not seen this since

Resolved as a tiny bump on a huge background
How do we reduce background:

\[ q \quad W \quad \bar{q} \quad \text{Jets} \]
What's the first goal?

Require $W$ to be a high $p_T$

Recoiling jet off $W$ gives the $W$ high $p_T$

Quarks decaying from high $p_T$ $W$ nearby each other

Resolve this as a single jet
What's the first goal?

Require $W$ to be a high $p_T$

By exploiting recoiling jet can trigger event.
What's the first goal?

Require $W$ to be a high $p_T$

Have not observed a single $W$ jet boson event from SM produce (Have only seen $W$ jets in top events)
Jet Substructure

Brief intro
A jet is a complicated object with many particles. Construct observables exhibit aggregate properties.
Jet grooming

- In order get a clean mass peak:
  - Adopt an approach that cleans excess radiation in a jet

\[ W \to qq \]

High \[ W \rho_T \]

Pileup/QCD radiation
Jet grooming

- In order get a clean mass peak:
  - Adopt an approach that cleans excess radiation in a jet

Jet grooming improves the resolution of jets

Pileup/QCD radiation
Jet substructure basics

Typically use two main properties

Likelihood jet is two pronged

Jet mass
Additional Substructure Features

- To reduce impact of pileup on substructure
  - Utilize the PUPPI algorithm

PUPPI gives best performance under PU conditions
ATLAS showed a W/Z peak

- ATLAS had an excess at 2 TeV

- Looked for a W/Z peak as a way validating perfor

<table>
<thead>
<tr>
<th>Events / 100 GeV</th>
</tr>
</thead>
<tbody>
<tr>
<td>10^4</td>
</tr>
<tr>
<td>10^3</td>
</tr>
<tr>
<td>10^2</td>
</tr>
<tr>
<td>10^1</td>
</tr>
<tr>
<td>10^0</td>
</tr>
</tbody>
</table>

- 580 < p_T < 670 GeV
- Splitting scale > 0.45
- Δy_{12} < 1.1

- No information except this plot
- Never published, specific cuts
Adding more information

- In 2016 ATLAS revealed plots with looser cuts
Finding the W/Z on a slope

How does the upward slope impact the search for a resonance?

Background is peaking after the bump helps to improve reduce background at W
What if we want to cut tighter?

Background is peaking after the bump helps to improve reduce background at W
A Quick Game
Let's play a game

- Let's try to inject a signal like this

Into a background 50 times larger than this
This signal looks very much like a W peak
Find the signal

- Can you see the signal?
Find the signal

- Can you see the signal?

Here is the signal at 125 GeV on a flat background
That was an easy one!
Find the signal

- Can you see the signal?

Here is the signal at 125 GeV on a flat background.
Now the background peaks just after the signal.
Find the signal

- Can you see the signal?

Here it is. This one is a bit harder.

Note exact same amount of background and signal

80 GeV
How about now?

- Can you see the signal?
• Can you see the signal?

Background
Note exact same amount of background and signal

How about now?

90 GeV
What's best background?

All variables yield roughly the same performance
Exactly the same background yield
Exactly the same signal yield

What if we could tune background shape to what we want?
What would be our best choice of shape?
Decorrelating Observables
Designing a new tagger?

To minimize sculpting we minimize the variation over mass and $p_T$.

$\tau'_{21} = \tau_2 / \tau_1 - M \times \rho'$

$\rho' = \rho + \log \left( \frac{p_T}{\mu} \right) = \log \left( \frac{m^2}{p_T \mu} \right)$
Does Designed tagger change perf?

- If transform done carefully *(raw is pre-transform)*:
  - No change or event improvement after transform

Post transform background is made flatter

Previously not flat
Does Designed tagger change perf?

- If transform done carefully (**raw is pre-transform**):
  - No change or event improvement after transform

Post transform background is made flatter

Post transform background
Performance has improved
Does Designed tagger change perf?

- If transform done carefully (raw is pre-transform):
  - No change or event improvement after transform

- Transformed variables preserve performance

Post transform background is made flatter

Post transform background Performance has improved
Does Designed tagger change perf?

- If transform done carefully (raw is pre-transform):
  - No change or event improvement after transform

We can now cut Really tight without sculpting our background

Post transform background is made flatter

Post transform background Performance has improved
After we get this new variable

• Finally decide to look for the W peak

• The strategy:
  
  – Select a jet:
    • Single jet with $p_T > 450$ GeV and (trimmed) mass $> 40$ GeV
  
  – Apply a substructure cut $\tau_2/\tau_1^{DDT}$
    • Cut really tight (~15-20% signal efficiency)
  
  – Veto on other objects in the jet
    • Electrons/muons/taus/Photons

  – Look for an excess in the groomed jet mass
Jet Mass distribution

CMS Preliminary

2.7 fb^{-1} (13 TeV)

- data
- QCD Pred.
- Total SM Pred.
- W(qq)
- Z(qq)
- Z'(qq), g_B = 1, m_Z = 200 GeV

W/Z \rightarrow qq
Jet Mass distribution

We see it, but
How exactly are we modeling the bkg
What's best background?

All variables yield roughly the same performance
Exactly the same background yield
Exactly the same signal yield

What if we could tune background shape to what we want?
What would be our best choice of shape?
What's best background?

All variables yield roughly the same performance
Exactly the same background yield
Exactly the same signal yield

Let me explain why

What if we could tune background shape to what we want?
What would be our best choice of shape?
One last ingredient

- Cutting on a variable gives us another advantage

A < Cut

Pass

Rich in signal

A > Cut

Fail

Poor in signal

We get two regions that we can study
One last ingredient

- Cutting on a variable gives us another advantage

A < Cut
Pass

A > Cut
Fail

Rich in signal

Poor in signal

We can merge these regions into one
Pass/Fail Ratio

- Combination of the two yield a single distribution

This gives us a new combined region
Pass/Fail Ratio

- Combination of the two yield a single distribution

We can fit this ratio instead of the mass distribution
How do you fit the ratio

- Fit both distributions at the same time

\[ \text{Bin}_{\text{Pass}}^i = \varepsilon \text{ Bin}_{\text{Fail}}^i \]
How do you fit the ratio

- Fit both distributions at the same time

If one bin moves up the other moves up
Even for weird shapes we fit a flat eff

$$\text{Bin}_{\text{Pass}}^i = \varepsilon \text{ Bin}_{\text{Fail}}^i$$
Forcing the ratio flat

- With a large MC sample we can make this flat.

Can construct a substructure observable to have a flat efficiency over mass.
In data

- We can tune the efficiency to be flat in MC

However, in data it might not be flat. We can fit for this deviation over mass.
In data

- We can tune the efficiency to be flat in MC

We can fit for this deviation over mass
Deviation from flat is a data/MC correction
Ingredients to see W peak

- Select a jet ($p_T > 500$ GeV)
- Require the jet to pass a 2 prong tag
  - Tune the tagger to be flat across the mass
- Select tagger & take both pass and failing selection
- Simultaneously fit for pass and failing
- Parametrize the difference between data and MC
  - Use the deviation from data and MC as the alternative

Before we go ahead and fit the mass

What exactly is the jet mass?
The Jet Mass
Anatomy of Quark/Gluon Jet Mass

Sudakov Peak

Fixed order

Resummed

Trend

 CMS, $L = 5$ fb$^{-1}$ at $\sqrt{s} = 7$ TeV, Ungroomed AK7 Dijets

$-\frac{1}{\sigma} \frac{d\sigma}{dm_J}$

$m_J^{AVG}$ (GeV)

Graph showing CMS data with different jet mass distributions and uncertainties.
Motion of the Sudakov Peak

By construction
Failing always has perfect agreement
Motion of the Sudakov Peak

Pass

Fail
Motion of the Sudakov Peak

![Graphs showing the motion of the Sudakov Peak](image-url)
Motion of the Sudakov Peak

Pass

Fail
Motion of the Sudakov Peak

Pass

Fail

$p_T$ 900-1000
Sudakov peak variation over $p_T$

For a fixed $p$ the features invariant over $p_T$

Jet Mass distribution varies with $p_T$

Aim to make flat mass & $p_T$
Anatomy of Jet Mass: Theory

Calculate in $\rho$ invariant over mass/$p_T$
Merging with $\rho$

- Merging to $\rho$ makes imposes invariance over $p_T$

When translating to $\rho$ distributions over $p_T$ are also invariant
This allows us to extend our fit from 1D mass to 2D $p_T$ and $\rho$
Design a transform to decorrelate against mass and $p_T$.

Decorrelating avoid mass sculpting allows us to cut tighter.
In data

- We can tune the efficiency to be flat in MC

Parametrize 2D efficiency by a surface in $\rho$ and $p_T$

$$N_{\text{pass}}^{QCD}(m_{SD}, p_T) = R_p/f(\rho, p_T) \times N_{\text{fail}}^{QCD}(m_{SD}, p_T)$$
In data

- We can tune the efficiency to be flat in MC

Utilize a polynomial in $\rho$ and $p_T$

$$N_{pass}^{QCD}(m_{SD}, p_{Tj}) = \epsilon^{QCD} \cdot \left( \sum_{k,\ell} a_{k\ell} \rho_{ij}^k p_{Tj}^\ell \right) \cdot N_{fail}^{QCD}(m_{SD}, p_{Tj}).$$
- Interatively expand in $\rho$ & $p_T$ order with an f-test

For W peak
Using
4$^{\text{th}}$ order in $\rho$
And
3$^{\text{rd}}$ order in $p_T$

\[ \rho = 2 \log(m_{SD}/p_T) \]

Utilize a polynomial in $\rho$ and $p_T$

\[ N_{\text{pass}}^{QCD}(m_{SDi}, p_{Tj}) = e^{QCD} \cdot \left( \sum_{k,\ell} a_{k\ell} \rho_{ij}^k p_{Tj}^\ell \right) \cdot N_{\text{fail}}^{QCD}(m_{SDi}, p_{Tj}) \]
Ingredients to see W peak: Recp

- Select a jet ($p_T > 500$)
- Require the jet to pass a 2 prong tag
  - Tune the tagger to be flat across the mass
- Select tagger & take both pass and failing selection
- Simultaneously fit for pass and failing
- Parametrize the difference between data and MC
- Use a 4$^{th}$, 3$^{rd}$ order polynomial in $\rho, p_T$ to model data

Now are we ready to fit the W peak?
The End Game
Meanwhile...

- Theorists decided to build in the scale invariance

**New Angles on Energy Correlation Functions**

Ian Moult, Lina Necib, Jesse Thaler

*(Submitted on 23 Sep 2016)*

- Into a new set of substructure observables
- Guiding principles are to exploit invariances in QCD
Full Decorrelation scheme

Use the k-Nearest Neighbor approach to determine $N_2^2$ cut

EXO-17-001
Clear W peak
Up to the last $p_T$ bin

At the highest $p_T$ other features are present

EWK Sudakov and NNLO W $p_T$ critical
Adding them all up

Sensitivity to the $W$ peak at the 10% level
Dark Matter: Recasting the $W/Z \rightarrow qq$ peak
What else?

- Without loss of generality we also have dijets

Mediator is coupling to quarks and to Dark matter
What else?

- Without loss of generality we also have dijets

This is a dijet+ISR search

Mediator is coupling to quarks and to Dark matter

Mediator can decay to quarks
Can use W peak to calibrate the signal strength across mass.
Combining Excess with 2.9 (2.2 local) significance
Visible in last bin

EXO-17-001
Low Mass $Z'$

CMS Preliminary

Axial-vector mediator

Dirac DM

$g_{DM} = 1.0$
$g_q = 0.25$
$g_i = 0$

$M_{Med} = 2 \times m_{DM}$

$\Omega_c h^2 \geq 0.12$

Exclusion at 95% CL

- Observed
- Expected

Dijet (35.9 fb$^{-1}$) [EXO-16-056]
Boosted dijet (35.9 fb$^{-1}$) [EXO-17-001]
DM + j/V(qq) (35.9 fb$^{-1}$) [EXO-16-048]
DM + $\gamma$ (12.9 fb$^{-1}$) [EXO-16-039]
DM + Z(ll) (35.9 fb$^{-1}$) [EXO-16-052]

*Original Plots by Tristan Du Pree

CMS*
Low Mass $Z'$

Direct Detection

$\sigma_{\text{SD-proton}}^{\text{DM}} [\text{cm}^2]$ vs $m_{\text{DM}} [\text{GeV}]$

CMS Preliminary

LHCP 2017

CMS observed exclusion 90% CL
Axial-vector med., Dirac DM; $q_i = 0.25$, $q_{\text{DM}} = 1.0$

- Boosted dijet (35.9 fb$^{-1}$) [EXO-17-001]
- Dijet (35.9 fb$^{-1}$) [EXO-16-056]
- $\text{DM} + j/V_{qq} (35.9 \text{ fb}^{-1})$ [EXO-16-048]
- $\text{DM} + \gamma (12.9 \text{ fb}^{-1})$ [EXO-16-039]
- $\text{DM} + Z_{\ell\ell} (35.9 \text{ fb}^{-1})$ [EXO-16-052]

DD/ID observed exclusion 90% CL

- PICASSO [arXiv:1611.01499]
- PICO-60 [arXiv:1702.07666]
- Super-K (bb) [arXiv:1503.04858]
- IceCube (bb) [arXiv:1612.05949]
- IceCube (tt) [arXiv:1601.00653]

*Original Plots by Tristan Du Pree

CMS-DP-2016-057
Hadronically Identified and Generated, after a Generation
Hadronically Identified and Generated, after a Generation

Searching for the Higgs
How do things change for Higgs

Require $W$ to be a high $p_T$

- Recoiling jet off $W$ gives the $W$ high $p_T$
- Quarks decaying from high $p_T$ $W$ nearby each other

Resolve this as a single jet
How do things change for Higgs

Require $W$ to be a high $p_T$

Initial state from gluons producing Higgs through a top loop

Jet remains two prongs
Also have 2 b-quarks
An Additional Check

As a check can look for $Z \rightarrow bb$

For a 500 GeV Boson:
Rate of $Z \rightarrow bb$ is about 10 times Higgs
Higgs Production

Largest production mode for Higgs is gluon fusion: Large background

Other modes have smaller rates & smaller backgrounds (you can tag something)
Ingredients to see H peak

- Select a jet $p_T > 450$ GeV (single massive jet trigger)
- Require the jet to pass a 2 prong tag
- Require jet to be double b-tagged
A b-tag?

Key to b-tagging is secondary vertex

Ripple effect on the resulting particles

typical jet from up quark
typical jet from bottom quark

“secondary” vertex where b hadron decayed
Concept #1

Apply standard b-tagging on the whole jet
Best at identifying Higgs against a jet composed of light quarks
Apply b-tagging on the subjet
Best at identifying Higgs against a jet from a gluon splitting

Concept #2

fat-jet b-tagging

sub-jet b-tagging

double-b tagger

Reject

$\tau$-axis$_1$

$\tau$-axis$_2$
Apply b-tagging on the Whole Jet using substructure Variables (n-subjettiness axes)
Double B-tagger

Resulting combination gives 50% improvement over previous
What about data/MC

Tag two muons in a jet
Use this to infer signal-like
2 b-quarks in a jet

Well modeled between data/MC
What about the correlation?

- Cutting on double b-tag is invariant of mass and $p_T$

No significant variation over mass

No significant variation over $p_T$

Not Correlated
Ingredients to see Higgs

- Select a jet $p_T > 450$ GeV (single massive jet trigger)
- Require the jet to pass a 2 prong tag
- Require jet to be double b-tagged
- Select tagger & take both pass and failing selection
Ingredients to see Higgs

- Select a jet $p_T > 450$ GeV (single massive jet trigger)
- Require the jet to pass a 2 prong tag
- Require jet to be double b-tagged
- Select tagger & take both pass and failing selection
Ingredients to see Higgs

- Select a jet $p_T > 450$ GeV (single massive jet trigger)
- Require the jet to pass a 2 prong tag
- Require jet to be double b-tagged
- Select tagger & take both pass and failing selection
- Simultaneously fit for pass and failing
Ingredients to see Higgs

- Select a jet $p_T > 450$ GeV (single massive jet trigger)
- Require the jet to pass a 2 prong tag
- Require jet to be double b-tagged
- Select tagger & take both pass and failing selection
- Simultaneously fit for pass and failing
- Parametrize the difference between data and MC
  - Use a $2^{\text{nd}}$, $1^{\text{st}}$ order polynomial in $\rho,p_T$ to model data
Ingredients to see Higgs

- Select a jet $p_T > 450$ GeV (single massive jet trigger)
- Require the jet to pass a 2 prong tag
- Require jet to be double b-tagged
- Select tagger & take both pass and failing selection
- Simultaneously fit for pass and failing
- Parametrize the difference between data and MC
  - Use a $2^{\text{th}}$, $1^{\text{st}}$ order polynomial in $p, p_T$ to model data

Now are we ready to fit the H peak?
The large $Z\rightarrow bb$ Allows us to calibrate our signal
The large $Z \rightarrow bb$ allows us to calibrate our signal.

Z peak
Allows us to calibrate Higgs.
Impact of Result

Probing gluon coupling vs top coupling

At high $p_T$, effective vertex dominated by top quark

Differences at high $p_T$:
Directly probe modifications in top quark coupling
Impact of Result

Many different parameterizations of modified gluon vs top couplings modification @ high $p_T$.
Observe a $Z \rightarrow \text{bb}$ with $5.1\sigma (5.9 \text{ exp}) \mu = 0.78^{+0.23}_{-0.19}$

Observe a $H \rightarrow \text{bb}$ with $1.5\sigma (0.7 \text{ exp}) \mu = 2.3^{+1.8}_{-1.6}$
Looking Forward
Data Flow in CMS

40 MHz

High speed
Low granularity readout (10μs)

100 kHz

Intermediate speed (100ms)
better readout

500 Hz

Full data readout (10s)

“Offline Computing”
Grid, O(10) Pb

Despite the large scale reduction we still store many
Petabytes of data
CMS Upgrade

40 MHz
Data reduction
5 kHz
Data reduction
1500 kHz

High speed
Low granularity
readout (10μs)

Intermediate speed (100ms)
better readout

Full data readout
(10s)

"Offline Computing"
Grid, O(10) Pb

This might become exabytes?!!!

See later
Upgrade L1 Trigger

New

Track Trigger
HGCAL Trigger

2.5 μs

1.0 μs

Muon

New

TRK
EC

EB TPG
HB TPG
HF TPG

EB
HB
HF

DT
RPC
CSC
GEM

BM TPG
RPC TPG
CSC TPG
GEM TPG

Barrel Calo Trigger
Barrel Muon Track Finder
Endcap Muon Track Finder

Correlator Trigger

Global Trigger

L1 Trigger Project
Visualizing what it does

- PUPPI: An image filter using QCD
  Reweights/removes all particles to eliminate pileup

To Perform this level of filtering on all collisions use custom electronics driven by a FPGA

TDR-17-004
Programming an FPGA

<table>
<thead>
<tr>
<th>CPU</th>
<th>GPU</th>
<th>FPGA</th>
</tr>
</thead>
<tbody>
<tr>
<td>Past</td>
<td>Present</td>
<td>Future</td>
</tr>
<tr>
<td>Speed: 1</td>
<td>Speed: 20-50</td>
<td>Speed: 200-1000</td>
</tr>
<tr>
<td>Single</td>
<td>Multiple simpler operations in parallel</td>
<td>Efficient packing of operations highly</td>
</tr>
<tr>
<td>complex</td>
<td>operation per clock</td>
<td>parallelized (low power)</td>
</tr>
<tr>
<td>operation</td>
<td>commonly available</td>
<td></td>
</tr>
</tbody>
</table>
Specialized code is needed to program the FPGA

Use a mixture of VHDL and a C based code called High Level Synthesis
What exactly are we doing?

Aim is to combine raw inputs into final state objects

Track Trigger + Calorimeter inputs = inside FPGA

Final output: combined particle flow physics objects

- Charged Hadron
- Neutral Hadron
- Electron
- Photon
- Muon
An example of FPGA programming

- Lets look at vertexing

The Fast

Fill N histograms with Δz in parallel

Scan and find max Δz value

Time: $\log_2(N^{tk})$ clocks + Scan
Size: $N^{tk} \times N^{\text{bins}} \times \log_2(N^{tk})$

The Slow

Scan and find max Δz value

Time: $N^{tk}$ clocks + Scan
Size: $N^{\text{bins}}$ (a single block RAM!)
Physics Benchmark

First performance studies show large improvements
We have only been developing this for on year

Added physics information for the same data taking rate

Select twice as many Higgs events with trigger increase

TDR-17-004
Beyond the Basics

- Can we identify the Higgs in the L1 trigger?

Higgs boson most commonly produced in a single jet of particles (jet) resulting from 2 b-quarks.

Complicated object relying on Machine Learning to indentify.
Arxiv Today!

Phil Harris (MIT),
Nhan Tran, Ben Kreis, Javier Duarte, Sergio Jindariani (FNAL),
Jen Ngadiuba, Maurizio Pierini (CERN)
Zhenbin Wu (UIC)
EJ Kreiner (Hawkeye 360)
Song Han (Stanford/Google→MIT)
High level synthesis: C-based compiler for FPGA

Machine learning

4 was once F
Can we put ML on an FPGA

- Neural Net inference how does it work?

\[
\ell_j^k = \phi(W_{ij}\ell_i^{k-1} + b_j)
\]

Activation function
Matrix mult
Vector addition
Can we put ML on an FPGA

- Neural Net inference how does it work?

\[ \ell^k_j = \phi(W_{i,j} \ell^{k-1}_i + b_j) \]
Can we put ML on an FPGA

- Neural Net inference how does it work?

\[ \ell_{j}^{k} = \phi(W_{ij}\ell_{i}^{k-1} + b_{j}) \]
A basic training design

- Start with a 3 layer Deep Neural Net
  - 16 high level features (not trigger friendly)
  - Jet mass and ECFs ($\beta=0,1,2$)
  - Output is a 5 dimensional discriminator

Test all the usual discriminators
Visualizing Input Layer 1

Layer 1
- 64 nodes
- Activation: ReLU

Layer 2
- 32 nodes
- Activation: ReLU

Layer 3
- 32 nodes
- Activation: ReLU

Softmax

- 16 inputs
- 64 nodes
- 32 nodes
- 32 nodes
- 5 outputs
- Activation: SoftMax

16*64 + 64*32 + 32*32 + 32*5 = 4,256 synapses
Overall Resources

Resources increase with Precision

3-layer pruned, Kintex Ultrascale

Beyond precision plateau
Precision Performance

Number of Multiplication units needed (typical chip has \( \sim 5000 \))

Number of clocks needed
5ns Clock size (200 MHz)

Have built one of the fastest neural nets in world
Summing Up the Data flow

Usual Training Step

- Model
- Keras, TensorFlow, PyTorch...
- Compressed model

hls4ml

- HLS conversion
- HLS project
- Co-processing kernel
- Custom firmware design

Tune configuration precision reuse/pipeline

Usual ML software workflow
Summing Up the Data flow

HLS4ML Step
Convert+Tune algo
Summing Up the Data flow

Embed Algo into custom firmware
What exactly is everybody doing?

Specialized Hardware for Machine Learning

#RealTimeAI
What exactly is everybody doing?

Many companies are investing in FPGA(-like) systems.
FPGAs have the 2nd best power performance consistent
ASICs better take years to program chip and its permanent
FPGAs can be reprogrammed in a matter of seconds

Outlook
Combining with the trigger
Combining with the trigger

New ISR types pushes limit to lower mass
Combining with the trigger and ML pushes limit down (more events)
Note with lifetime added
Relic bound is roughly the same
Conclusion

• We have developed:
  – A new approach to look for low mass resonances
  – Identify them as a hadronic jet

• In a single jet we have shown:
  – First observation of the $W \rightarrow qq$
  – First observation of the $Z \rightarrow bb$
  – New and interesting bounds on low mass mediators
  – Proof that we can be sensitive to high $p_T$ Higgs

• We are still early on in the development
  – Adding the trigger will really help to push this further
Here is where we started

Phil Harris (MIT)
With these new tools
Opening a new era of precision

Phil Harris
(MIT)
Thanks!
What exactly are we doing?

Correlator Layer 1

Final output are Pileup cleaned Particles

Currently working on algos for high-level objects

Physics Objects

Jets

MET

Electron Photon

Muons

Demonstrated algorithm on boards @CERN/FNAL

Correlator Layer 1
What exactly are we doing?

Correlator Layer 1

- Single Board
  - Tracks
  - Calo Clusters
  - EM Clusters
  - Muon

Linking

- Mu Tk
- EM Tk
- Ca Tk

Vertexing

- PUPPI Cands

Final output are Pileup cleaned Particles

Currently working on algos for high-level objects

Layer 2

- Physics Objects
  - Jets
  - MET
  - Electron Photon
  - Muons
CMS Upgrade

First chain in readout is the “Level 1 Trigger”
MIT(PH) is working on the upgrade

High speed
Low granularity readout (10μs)

Intermediate speed (100ms)
Better readout

Full data readout (10s)
Ingredients to see H peak

• Select a jet
  - $p_T > 450$ GeV (single jet trigger w/trimmed mass $> 40$ GeV)
Ingredients to see H peak

- Select a jet: $p_T > 450$ GeV (single massive jet trigger)
- Require the jet to pass a 2 prong tag
  - $N_2^{DDT}$ at the 25% bkg efficiency
  - This time we will not fit for the efficiency of $N_2$
Uses of substructure

- Substructure allows us to reduce background

Cutting on Mass & substructure reduces background by 97-98% (50% sig eff)

Cutting on Mass alone
……at the same time on CMS

• An observation was being made

Background with current cuts force an upward slope
Adding Dark Matter

- What drives dark matter interaction is production
  - Take the approach that this is defined by the mediator

\[ \mathcal{L}' = g_{\text{DM}} \chi \chi Y \]

- \( Z'^\mu \) Spin 1
  - Uniform coupling to SM
  - \[ \mathcal{L}' = \mathcal{L}' + g_{\text{SM}} Z'^\mu \bar{q} \gamma^\mu q \]

- \( S \) Spin 0
  - Yukawa* couplings to SM
  - \[ \mathcal{L}' = \mathcal{L}' + g_{\text{SM}} S \bar{q} q \]
Dark Matter gives 2 types

- Couplings to SM force two different scenarios

Spin 0

Yukawa coupling to quarks
Dominated by heavy quarks

Small cross sections:
Probe low masses or
Large couplings

Spin 1

Flavor universal to quarks
Dominated by light quark

Large cross sections:
Probe large masses or
Small couplings
New Substructure Observables

Using AK8 PUPPI jets

We use $N_2$ to see the W peak
We fit over a fixed $p_T$ range yields disjoint mass.
Note with lifetime added
Relic bound is roughly the same

Coupling Bound

Relic bound below we overclose
(see https://arxiv.org/abs/1508.03050)
Note with lifetime added
Relic bound is roughly the same

Relic bound below we overclose
(see https://arxiv.org/abs/1508.03050)
Coupling Bound

Boosted Dijet Projected

Vector Projection $3ab^{-1}$

$\frac{m_{DM}}{m_{Med}} = \frac{1}{3}$

$m_{DM}(GeV)$

$10^3$
For masses down to 10, exclude all couplings that give relic.
Vector Coupling Bound

Vector Projection $3a b^{-1}$

$m_{\text{DM}}/m_{\text{Med}} = \frac{1}{3}$

$m_{\text{DM}} (\text{GeV})$ vs $g_\rho$

- Monojet best limit
- MA best limit
- TK best limit
- $Z \rightarrow qq+j$
- $Z \rightarrow qq$
- $\Omega h^2$
Current Best Higgs $p_T$ measurement

CMS Preliminary 35.9 fb\(^{-1}\) (13 TeV)

No other way to look for events in this region
Previously this has been done by two experiments

CDF observed small $b\bar{b}$ excess on a large background

ATLAS observed $Z \rightarrow b\bar{b}$ at lower $p_T$

Both had small sensitivity to Higgs: deemed impossible
The Higgs boson at high $p_T$

Tobias Neumann* and Ciaran Williams†

Nearly every paper stops at 400 GeV or less

Higher $p_T$ is thought to be impossible
• No experiment has quoted $3\sigma$ sensitivity
  - Combined ATLAS/CMS is $2.7\sigma$ observed $3.7\sigma$ expected

$H\rightarrow bb$ strength is low and not well confirmed

No one has yet to consider the gluon fusion channel
For H peak

2^{nd} order in \rho
1^{st} order in p_T

Following

\[ N_{\text{pass}}^{QCD}(m_{SDi}, p_{Tj}) = e^{QCD} \cdot \left( \sum_{k,\ell} a_{k\ell} \rho_{ij}^k p_{Tj}^\ell \right) \cdot N_{\text{fail}}^{QCD}(m_{SDi}, p_{Tj}) \]
3 Elements to the Best Higgs $p_T$

#1 : Finite top quark mass in loop known to LO

#2 : Higher order corrections with approx NLO
   Finite Top mass

#3 : Higher order corrections with Infinite Top mass (NNLO)
3 Elements to the Best Higgs $p_T$

#1: Finite top quark mass in loop known to LO

#2: Higher order corr With approx NLO Finite Top mass

#3: Higher order corrections with Infinite Top mass (NNLO)

Central prediction is a hybrid of 3 approaches

$$G F H(\text{NNLO} + m_t) = \text{Powheg}(1 \text{ jet } m_t \to \infty) \times \frac{\text{MG LO } 0-2 \text{ jet } m_t}{\text{Powheg}(1 \text{ jet } m_t \to \infty)} \times \frac{\text{NLO } 1 \text{ jet } m_t}{\text{LO } 1 \text{ jet } m_t} \times \frac{\text{NNLO } 1 \text{ jet } m_t \to \infty}{\text{NLO } 1 \text{ jet } m_t \to \infty}.$$
Hints?

**ATLAS and CMS**

**LHC** Run 1

- **Top High**
- **Gluons Low**

- $|\kappa_V| \leq 1$
- $B_{BSM} \geq 0$
- $B_{BSM} = 0$

Parameter value
Dark matter benchmark
Dark matter benchmark
Dark matter benchmark
Dear Authors

I noticed that in your PAS the introduction refers to UA1 and UA2 results at sqrt(s)=300 GeV. These results came out when I was on UA2 in fact, and the SppS ran at sqrt(s)=630 GeV.

Best regards Joe[Incandela]
Broke a record almost 30 years old

Incidentally this did not happen at the olympics
Probing the Mass range

Like Monojet, we can expand to further regions by tagging other objects.

Standard jet triggers

No tag
Probing the Mass range

Trigger Level analysis

Jets in trigger

Standard jet triggers

450-2000 GeV

1100-8000 GeV
Probing the Mass range

Jet ($p_T > 430$ GeV) + 2-jets

ATLAS Preliminary

- $X + |P_T| < 0.6$
- $|\eta_j| < 0.6$

$|s| = 13$ TeV, 15.5 fb$^{-1}$

- Data
- Fit
- Background fit
- BumpHunter interval

300-600 GeV

450-2000 GeV

1100-8000 GeV

jj+j ISR analysis

Trigger Level analysis

Standard jet triggers

Additional jet (like monojet) ATLAS
Probing the Mass range

**γ (p_T>150 GeV) + 2-jets**

*ATLAS Preliminary*

- Events
- $X + \gamma (p_T > 150 \text{ GeV})$
- $|\eta_{\gamma}| < 0.8$
- ATLAS Preliminary
- $\sqrt{s} = 13 \text{ TeV}$, $15.5 \text{ fb}^{-1}$
- Data
- Background fit
- BumpHunter interval
- 200-1400 GeV

**Jet (p_T>430 GeV) + 2-jets**

*ATLAS Preliminary*

- Events
- $X + j (p_T > 430 \text{ GeV})$
- $|\eta_j| < 0.8$
- ATLAS Preliminary
- $\sqrt{s} = 13 \text{ TeV}$, $15.5 \text{ fb}^{-1}$
- Data
- Background fit
- BumpHunter interval
- 300-600 GeV

**Trigger Level analysis**

*CMS Preliminary*

- Events
- $X + j (p_T > 430 \text{ GeV})$
- $|\eta_j| < 0.8$
- CMS Preliminary
- $\sqrt{s} = 13 \text{ TeV}$, $12.9 \text{ fb}^{-1}$
- Data
- Fit
- $gg (750 \text{ GeV})$
- $gg (1200 \text{ GeV})$
- $gg (1600 \text{ GeV})$
- Fit Range: 303-611 GeV
- 450-2000 GeV

**Standard jet triggers**

*CMS Preliminary*

- Events
- $X + j (p_T > 430 \text{ GeV})$
- $|\eta_j| < 0.8$
- CMS Preliminary
- $\sqrt{s} = 13 \text{ TeV}$, $12.9 \text{ fb}^{-1}$
- Data
- Fit
- $gg (2.0 \text{ TeV})$
- $gg (4.0 \text{ TeV})$
- $gg (6.0 \text{ TeV})$
- Fit Range: 1100-8000 GeV

**Additional photon (like monophoton)**

*ATLAS*

**EXO-16-030**
Many people can get deceived by data in front of them.
That's what we were looking at.

There is much more we can do with the LHC.
We are just learning the tools now.
Dijet A(Pseudoscalar) Z' Scalar Low mass Empire
After a 25 year hiatus, rekindling the standard model with low mass hadronic decays

Phil Harris (CERN)
Thanks!
New Directions

- Improved b-tagging

Potential large improvements in b-tagging
From new layer & more inclusive approaches
New Directions

- Key concept was modifying jet substructure

New approaches w/deep learning can decorrelate observables
Large Modifications at high $p_T$
How does this translate to DM?

With the standard cross section formula

\[ \sigma = \frac{g_{\text{SM}}^4 C}{\Gamma M_{\text{med}}^4} \]

Dijet cross section@LO

\[ \sigma_0(g_B) = \sigma_{\text{DM}}(g'_B, g_{DM} = 1, m_{DM}) \]

Add DM contribution

Coupling bound

\[ (g'_B)^2 = \frac{g_B^2}{2} \left( 1 + \sqrt{1 + 4 \frac{\Gamma_{DM}}{\Gamma(g_B)}} \right) \]

Solve for coupling bound with DM

\[ m_{\text{DM}}(\text{GeV}) \]

\[ m_{\text{med}}(\text{GeV}) \]
Question:

- What's the simplest way to present LHC results in the context of Dark Matter?
Question:

- What's the simplest way to present LHC results in the context of Dark Matter?

Answer:

- $\sigma_{\text{Invisible}}$

- Assumes dark matter coupling to standard model

- $\mathcal{L}' = g_{\text{DM}} \chi \bar{\chi} Y$ + SM interactions

Diagram:

- Dark Matter
- Mediator
- + SM interactions
Adding Dark Matter

- What drives dark matter interaction is production
  - Take the approach that this is defined by the mediator

$$\mathcal{L}' = g_{\text{DM}} \chi \chi Y$$

- **Z'** Spin 1
  - Uniform coupling to SM
  $$\mathcal{L}' = \mathcal{L} + g_{\text{SM}} Z'_\mu \bar{q} \gamma^\mu q$$

- **S** Spin 0
  - Yukawa* couplings to SM
  $$\mathcal{L}' = \mathcal{L} + g_{\text{SM}} S \bar{q} q$$
Preserving Generality?

To compare with other (low energy) searches:
split by spin dependence

\[
\mathcal{L}' = g_{\text{DM}} \chi \bar{\chi} Y
\]

\[
\mathcal{L}' = \mathcal{L}' + g_{\text{SM}} Z' \bar{q} \gamma^\mu \gamma^5 q
\]

\[
\mathcal{L}' = \mathcal{L}' + g_{\text{SM}} Z' \bar{q} \gamma^\mu \gamma^5 q
\]

\[
\mathcal{L}' = \mathcal{L}' + g_{\text{SM}} S \bar{q} q
\]

\[
\mathcal{L}' = \mathcal{L}' + g_{\text{SM}} P \bar{q} \gamma^5 q
\]

Strategy of searches in LHC does not change much
Interpretation against Direct Detection/Indirect Changes a lot
Simplified Models 101

**Vector**

\[ g_{DM} Z'_\mu \bar{\chi} \gamma^\mu \chi \]

EWK style coupling
(equal to all leptons)

**Axial**

\[ g_{DM} Z''_\mu \bar{\chi} \gamma^\mu \gamma^5 \chi \]

EWK style coupling
(equal to all leptons)

**Scalar**

\[ g_{DM} S \bar{\chi} \chi \]

Yukawa style coupling
(Mass based coupling)

**Pseudoscalar**

\[ g_{DM} P \bar{\chi} \gamma^5 \chi \]

Yukawa style coupling
(Mass based coupling)
Relic Density
What's the form of the Relic?

- Considering the form
Relic Density on Simplified Models

- Relic constraints obtained using madgraph models
  - Calculation is performed with MadDM (1.0)
    - Well validated program can be used with madgraph
    - For models use LO simplified models
  - Models obtained through DMF exercise
Heuristic example to how it works

1. $m_\chi < m_t \rightarrow$ tight constraints
   - Suppression of the annihilation
   Process $\chi \chi \rightarrow \text{Med} \rightarrow \text{SM}$

2. $m_\Phi \sim 2 \cdot m_\chi \rightarrow$ weaker constraints
   - On-shell $\chi \chi \rightarrow \text{Med}$

3. $2 \cdot m_\chi > m_\Phi \rightarrow$ weaker constraints
   - Less suppressed annihilation $\chi \chi \rightarrow \text{Med} \rightarrow \text{SM}$
How does it look for our 4 friends?

Coupling is set to 1 for all plots

Vector

Axial

Scalar

Pseudo
How does it look for our 4 friends?

Coupling is set to 1 for all plots

- **Vector**
  - Max at 70 TeV

- **Axial**
  - Max at 9 TeV

- **Scalar**
  - Max at 7 TeV

- **Pseudo**
  - Max at 35 TeV
The Basic Monojet Search

Escaping detector gives us signatures of $MET$
Searching for $MET$

$-\sum_{\text{All particles}} p_T = MET_{(E_T^{\text{Miss}})}$

$-\text{Boson } p_T = MET_{(E_T^{\text{Miss}})}$

“To find nothing you have to reconstruct everything”[1]
How do we search?

$Z \rightarrow \nu \nu$

Just the $Z$ $p_T$ spectra

Signal

$E_T^{\text{miss}}$ (GeV)

19.7 fb$^{-1}$ (8 TeV)
How do we search(data)?

MonoJet Selection
How do we search(data)?

MonoJet Selection

Its a precision analysis
How do we get to this precision?

Rely on a series of control regions to correct for the data/MC agreement
Strategy to fix agreement

Control Region

Data

Control: another decay of a Z boson

\( \text{Z} \rightarrow \mu \mu \)  \( \times \) Remove

\( \text{Z} \rightarrow \nu \nu \)

hadronic recoil : Transverse sum of all particles in event excluding leptons/photons

Signal Region

Data

Propagate scale factor from a control region w/similar \( p_T \)

MC
What is the transfer factor?

Propagate the data/MC agreement of the hadronic recoil
From a control region to a signal region
What is the transfer factor?

Do this formally by embedding transfer factors $R_i$

Easily propagate these into fitted likelihood for signal
Control regions have less events than signal

\[ \sigma_{\mu\mu} = 0.1 \sigma_{\nu\nu} \]

Statistical precision is 4x worse
Not good enough!
2 Control regions
60% uncertainty @ 1 TeV

CMS-EXO-16-052
3 Control regions
40% uncertainty @ 1 TeV
4 Control regions
30% uncertainty @ 1 TeV

CMS-EXO-16-052
5 Control regions
15% uncertainty @ 1 TeV

CMS Preliminary

Z → μμ
W → ev
W → μν
Z → ee
γ + jets
5 Control regions + Signal
15% uncertainty @ 1 TeV
All in one big Simultaneous fit

CMS-EXO-16-052
The Partial Likelihood of the fit

\[
\mathcal{L}(\mu, \mu^Z, \theta) = \prod_i \text{Poisson} \left( d_i^\gamma | B_i^\gamma(\theta) + \frac{\mu_i^Z}{R_i^\gamma(\theta)} \right) \\
\times \prod_i \text{Poisson} \left( d_i^Z | B_i^Z(\theta) + \frac{\mu_i^Z}{R_i^Z(\theta)} \right) \\
\times \prod_i \text{Poisson} \left( d_i^W | B_i^W(\theta) + \frac{f_i(\theta) \mu_i^Z}{R_i^W(\theta)} \right) \\
\times \prod_i \text{Poisson} \left( d_i | B_i(\theta) + (1 + f_i(\theta)) \mu_i^Z + \mu S_i(\theta) \right)
\]

\( d_i \) = Number of events

\( B_i(\theta) \) = MC Prediction

\( f_i(\theta) = \frac{\mu^W}{\mu^Z} \)

**γ+jets**

**Z+jets**

**W+jets**

 Signal Region

Addition of signal

Systematics are not included
Can we really use all these regions?
Answer:
Sort of.....
A mystery? Understanding $Z/\gamma p_T$

How are going to use photons for $Z$ with this kind of prediction?

We need this to model the ratio perfectly to use it
2015
Impact of the electroweak corrections

How do we fix this?

We care about the of the two
Solving the Mystery

Adding the EWK corrections brings back agreement

Some of the more heinous diagrams

CMS $Z/\gamma$ ratio (8 TeV) measurement compared to different generators
Solving the Mystery

Some of the more heinous diagrams

Adding the EWK corrections brings back agreement
2014...again!
However we still have a problem!

Unc. \[ \frac{d\sigma^Y}{dp_T} / \frac{d\sigma^Z}{dp_T} = \frac{d\sigma^Y}{d\sigma^Z(\mu)} \]

We don't know how to do scale uncertainties on ratios!
However we still have a problem!

\[ \text{Unc.} \rightarrow \frac{d\sigma^Y}{dp_T} \bigg/ \frac{d\sigma^Z}{dp_T} = \frac{d\sigma^Y}{d\sigma^Z}(\mu) \]

We don't know how to do scale uncertainties on ratios!
Uncertainty on ratio? How is it done?

Scale uncertainty on process #1
Scale uncertainty on process #2
Uncertainty on process #1/process #2 (fully correlated)

\[
\frac{d\sigma^Y}{dp_T} \bigg/ \frac{d\sigma^Z}{dp_T} = \frac{d\sigma^Y}{d\sigma^Z(\mu)}
\]

\[
\begin{pmatrix}
  d\sigma^Y(+\sigma) \\
  d\sigma^Z(+\sigma)
\end{pmatrix} =
\begin{pmatrix}
  1 & 0 \\
  0 & 1
\end{pmatrix}
\begin{pmatrix}
  d\sigma^Y(\mu^\text{up})/d\sigma^i(\mu_0) \\
  d\sigma^Z(\mu^\text{up})/d\sigma^i(\mu_0)
\end{pmatrix}
\]
Uncertainty on ratio? How is it done?

Scale uncertainty on process #1
Scale uncertainty on process #2
Uncertainty on process #1/process #2 (fully correlated)

Adjust $C$ until uncertainty is
Uncertainty on ratio? How is it done?

Scale uncertainty on process #1
Scale uncertainty on process #2
Uncertainty on process #1/process #2 (fully correlated)

Unc. \( \frac{d\sigma_Y}{dp_T} / \frac{d\sigma_Z}{dp_T} = \frac{d\sigma_Y}{d\sigma_Z}(\mu) \)

\[
\begin{pmatrix}
  d\sigma_Y(+\sigma) \\
  d\sigma_Z(+\sigma)
\end{pmatrix} =
\begin{pmatrix}
  1 & C \\
  C & 1
\end{pmatrix}
\begin{pmatrix}
  d\sigma_Y(\mu_{up})/d\sigma_i(\mu_0) \\
  d\sigma_Z(\mu_{up})/d\sigma_i(\mu_0)
\end{pmatrix}
\]

Decorrelate scale unc. until its max of either process

\( d\sigma_Y/d\sigma_Z (+\sigma) < \max_i (d\sigma_i(\mu_{up})/d\sigma_i(\mu_0)) \)
What is the previous unc?

Can we motivate this?

\[
\text{Unc.} \quad \frac{d\sigma^y}{dp_T} \bigg/ \frac{d\sigma^z}{dp_T} = \frac{d\sigma^y}{d\sigma^z}(\mu)
\]

\[
\begin{pmatrix}
  d\sigma^y(+\sigma) \\
  d\sigma^z(+\sigma)
\end{pmatrix} =
\begin{pmatrix}
  1 & C \\
  C & 1
\end{pmatrix}
\begin{pmatrix}
  d\sigma^y(\mu^{up})/d\sigma^i(\mu_0) \\
  d\sigma^z(\mu^{up})/d\sigma^i(\mu_0)
\end{pmatrix}
\]

Makes Little Sense
What about EWK corr. uncertainty?

In light of being conservative:
Take the full correction
What about EWK corr. uncertainty?

In light of being conservative:
Take the full correction

Additionally de-correlated this per bin
Avoids low MET to high MET constraints

We were forced to do this by management!
This makes us too conservative
2015...
We are stuck
How do we propagate this in?

The actual uncertainties

Correlated across boson $p_T$

EWK: Since its not correlated dominant in tails
How exactly do they look?

- Z/γ Ratio
- Z/W Ratio
What do the uncertainties look like?

Updated unc still too large
Profiling them in the fit

Constraints after the fit

Still systematics limited @low MET
Not systematics limited @ high MET
→ Likely will never be
Before fitting

After fitting
2016:
A new hope
Its clear this is a major issue!

- Theorists now working on new recommend
  - MLM/S. Dittmaer/Pozzarin/Lindert/G. Salam
  - Still a work in progress

Delivered on Dec. 23 2015
What can we expect?

- Large reduction in theory uncertainties
  - Experimental effects of the same order

- Can now correlate unc. across boson $p_T$ bins

Reducing these unc. by a 5x can give x2 improvement (@ low mass)
Current Monojet Sensitivity

CMS Preliminary

monojet

Events / GeV

$35.9 \text{ fb}^{-1} (13 \text{ TeV})$

Data
Z(\nu \bar{\nu})+jets
WW/WZ/ZZ
Top quark
Z/\gamma(\gamma)+jets
QCD
Higgs invisible, $m_H = 125 \text{ GeV}$
Axial-vector, $m_{\text{med}} = 2.0 \text{ TeV}$

$E_T^{\text{miss}} [\text{GeV}]$

$E_T [\text{GeV}]$

$M_{\text{med}} = 500$
$M_{\text{med}} = 1000$
$M_{\text{med}} = 2000$

$q \rightarrow q X$
Dark Matter Mass

Pick a Model

CMS Preliminary

Vector med, Dirac DM, \( g_q = 0.25 \), \( g_{DM} = 1 \)

- - - Median expected 95\% CL
- - - 68\% expected

Observed 95\% CL

Observed = theory unc.

\( \Omega_c x h^2 \approx 0.12 \)

Observed \( \sigma_{95\%CL}/\sigma_{th} \)

Mass of mediator produced
Understanding sensitivity

CMS Preliminary

Vector med, Dirac DM, $g_q = 0.25$, $g_{DM} = 1$

- - - Median expected 95% CL
-- - 68% expected

- - - Observed 95% CL
- - - Observed = theory unc.

$\Omega_c x h^2 \approx 0.12$

Region we exclude
Dark Matter Models
SMS 101: monojet final state in collider

- **Vector (Spin independent)**
- **Axial (Spin dependent)**
- **Scalar (Spin independent)**
- **Pseudoscalar**
With Collider

Vector(SI)
Large cross section

Axial (SD)
Same as vector

Scalar(SI)
Low-ish cross section

Pseudoscalar
Better than scalar
Small cross section
<table>
<thead>
<tr>
<th></th>
<th>With Direct Detection</th>
</tr>
</thead>
<tbody>
<tr>
<td>Vector(SI)</td>
<td>So-so</td>
</tr>
<tr>
<td>Spin independent</td>
<td>Extremely good</td>
</tr>
<tr>
<td>Scalar(SI)</td>
<td>Forget about it*</td>
</tr>
<tr>
<td>So-so</td>
<td>Spin independent</td>
</tr>
</tbody>
</table>

With Direct Detection

Vector (SI)
Spin independent
Extremely good

Axial (SD)
Spin dependent
Not so great

Scalar (SI)
So-so
Spin independent

Pseudoscalar

Use indirect detection
With Direct Detection

- Vector (SI)
  - Spin independent
  - Extremely good
- Scalar (SI)
  - So-so
  - Spin dependent
- Axial (SD)
  - Spin dependent
  - Not so great
- Pseudoscalar
  - Use indirect detection
Now that search is cast in terms of mediator
No concerns in the translation

$$\sigma_{SI} = \frac{f^2 (g_q) g_{DM}^2 \mu_{nX}^2}{\pi M_i^A}$$
With Direct Detection

**CMS Preliminary**

**Vector(SI)**

- CMS exp. 90% CL
- CMS obs. 90% CL
- LUX
- CDMSLite
- PandaX-II
- CRESST-II

Scalar(SI)

- CMS obs. 90% CL
- LUX
- CDMSLite
- PandaX-II
- CRESST-II

**Axial(SD)**

- CMS exp. 90% CL
- CMS obs. 90% CL
- PICO-60
- Picasso
- IceCube bb
- IceCube tt
- Super-K bb

**Pseudoscalar**

- CMS exp. 90% CL
- CMS obs. 90% CL
- FermiLAT
Beyond Monojet
How Do We Discriminate Models?

Mono-jet

Models: Vector, Axial, Scalar, Pseudoscalar
How Do We Discriminate Models?

Mono-Photon

Models: Vector, Axial

A Photon

\[ q \rightarrow \phi \rightarrow x \]

Escapes detector

MET
How Do We Discriminate Models?

Mono-W

Monos : Vector, Axial, Higgs

Again Escapes detector

MET
Spin 1
Spin 1 DM Searches

Spin 1 production on SM couplings for final state
Easily extend this to other final states
Spin 1 DM Searches

Can look for a Vector boson+$MET$ as well

Replace w/Boson

$V \rightarrow qq/ll$
The split in simplified model terms

- With spin 1 can generate other final states:

\[ V \rightarrow qq \]

This is just a monojet
Spin 1 DM Searches

Can look for a Photon+MET as well

Replace w/Boson

Replace w/Photon

V→qq/ll
Spin 1 DM Searches

If vertex is flavor changing

If vertex is flavor changing

Replace w/Boson

Replace w/Photon

Replace w/top

Flavor changing vertex

top
The split in simplified model terms

- With spin 1 can generate other final states:
  
  This is just a monojet

Flavor changing vertex

Replace w/ top

Top
Vector Mediator

Collider Constraints
Vector Mediator

Collider Constraints

Direct detection constraints
Vector Mediator

Collider Constraints

Direct detection constraints

Limit of sensitivity of direct detection
Axial Mediator
Beyond Invisible Searches