Invisible Higgs & trigger challenges on ATLAS

A discussion on Dark sector, Higgs boson, Trigger, and ML on FPGA

Tae Min Hong

Université de Genève
September 28, 2021
Outline

Motivation (10 min)

- Dark sector
- Higgs boson

Analyses (20 min)

- $H_{125} \rightarrow$ Dark matter pair
  - ATLAS Collab, ATLAS-CONF-2020-008 (2020)
- $H_{125} \rightarrow$ Dark photon + $\gamma$
  - ATLAS Collab, ATLAS-CONF-2021-004 (2021)

Trigger (30 min)

- $E_T^{\text{miss}}$, VBF
  - ATLAS, JHEP 08 (2020) 080
- ML on FPGA
  - Hong et al., JINST 16 (2021) P08016

Higgs portal to DM? How to trigger?
Motivation

Dark sector
Higgs boson
Many evidence of dark matter
Here: colliding galaxy clusters

3 MLY (MACSJ0025, 2008)

Inferred distribution of matter by lensing

- X-rays from known matter
- Dark matter (Known unknown)

Source: https://apod.nasa.gov/apod/ap080917.html

Can we create dark matter? Related to Higgs?
Higgs boson couplings
Higgs couples to everything

Table of elementary particles

Quarks

<table>
<thead>
<tr>
<th>u</th>
<th>c</th>
<th>t</th>
</tr>
</thead>
<tbody>
<tr>
<td>d</td>
<td>s</td>
<td>b</td>
</tr>
</tbody>
</table>

Leptons

<table>
<thead>
<tr>
<th>e</th>
<th>μ</th>
<th>τ</th>
</tr>
</thead>
<tbody>
<tr>
<td>ν_ν_</td>
<td>ν_μ</td>
<td>ν_τ</td>
</tr>
</tbody>
</table>

Forces

<table>
<thead>
<tr>
<th>Z</th>
<th>γ</th>
</tr>
</thead>
<tbody>
<tr>
<td>W</td>
<td>g</td>
</tr>
</tbody>
</table>

spin-½ spin-0 spin-1
fermions bosons

Higgs coupling to each particle is determined given $m_H$

Source: Fermilab
Higgs is so narrow because 125 GeV

Higgs width in MeV (for $m_H = 125$)

Tree relation for massive

Loop relation for massless

Source: CERN Yellow Report (2014)

Massive bosons  p.s. suppressed ($m_H \ll 2 \cdot m_{w,z}$)

Fermion $m_F$  means tiny Yukawa couplings ($tt$ large, but $m_H \ll 2 \cdot m_{top}$)

Massless bosons  loop suppressed

Everything is suppressed
O(MeV) is not unreasonable
Portals to NP can look like

\[ \Gamma \]

Measure these branching ratios to indirectly limit invisible

Source: CERN Yellow Report (2014)

Theory
- SM singlet \( \sim g H^2 \chi^2 \)
- Fully renormalizable

Goncalves, Han, Mukhopadhyay, PRL 120, 11801 (2018)
Curtin +12 others, 1312.4992 (2014)
Chang +3 others, 0801.4554 (2008)
Silveira & Zee, PL B161,136 (1985)
and many many more papers.

Coupling at 0.01 \( \rightarrow \) MeV-level modification \( \rightarrow \) Large rate
What's allowed by individual measurements?

Measure each predicted slice

Constrains non-Standard Model (with caveats)

A fraction of Higgs decays *could* be related to our hypothesis

Source: [http://cdsweb.cern.ch/record/2629412](http://cdsweb.cern.ch/record/2629412)
**Dark matter** \((m>0)\)

**Related production channels**

---

**Dark matter**

\[
\begin{align*}
q & \rightarrow H \, q \, q \\
\chi & \rightarrow E_{T^{\text{miss}}} \, q \\
\end{align*}
\]

**Higgs production**

- ggF: No observable!
- ggF + 1 jet: Overwhelming strong bkg'd
- VBF: Depends on trigger threshold
- ZH: Suppressed by \(\sigma \cdot B\)
- WH: " and neutrino / hadronic W

**Dark matter + \(\gamma\)**

\[
\begin{align*}
q & \rightarrow H \, q \, \chi \\
\gamma & \rightarrow E_{T^{\text{miss}}} \, \text{photon} \\
\end{align*}
\]

**Comparison to without \(\gamma\)**

- Smaller signal size, but clean
- Adding \(\gamma\) reduces strong background
- Depends on trigger threshold
- Added bonus in interpretation (next slide)

**Trigger threshold is critical aspect of the study**
Dark matter \((m>0) \rightarrow\) Dark photon \((m=0)\)

Maybe there is a Dark sector

**Theory**

- Unbroken \(U(1)_{\text{dark}}\) with enhanced \(H \rightarrow \gamma \gamma_{\text{dark}}\)
- Signal peaks in the \(m_{T}\) of \((E_{T}^{\text{miss}}, \gamma)\) system

Expand the scope of the search with alternate signal models
Analyses

$H_{125} \rightarrow \text{Dark matter pair}$

$H_{125} \rightarrow \text{Dark photon} + \gamma$
Physics of VBF $H \rightarrow \text{invisible}$

VBF production of the Higgs is established

- Energetic jets with large $\eta$ gap
- No hadronic activity
- $m_{jj}$, $\Delta \eta_{jj}$, $N_{\text{central jets}}$

---

Fig. 1. Higgs boson production from virtual vector boson pairs ($V = W$ or $Z$).

Source: Cahn, Dawson, PLB 136 (1984) 196

---

Not many handles, background estimation is crucial
Signal & background

Z & W are largest

Signal

“Irreducible”

Lost lepton

Z + 2j via strong process

W + 2j via strong process

Z + 2j via weak process

W + 2j via weak process

Strong process is dominant, but weak process imp't at higher $m_{jj}$
Event display
High $E_T^{\text{miss}}$ (564 GeV) + High $m_{jj}$ (3.6 TeV)

High $E_T^{\text{miss}}$ event + two energetic jets with large $\Delta\eta$
Angular characteristics
Data distribution of separation in $\eta$

ATLAS, J. High Energy Phys. 01 (2016) 172

Use kinematic properties to statistically separate samples
Selection
Signal region, control regions

Cuts

- Two jets > 80 GeV, 50 GeV
- Centrality of additional jets
- $E_T^{\text{miss}} > 200$ GeV

Signal region

- Bin in $m_{jj} \otimes N_{\text{jet}} \otimes \Delta \phi_{jj} = 5 \otimes 2 \otimes 2$

Control region

- $W \rightarrow \ell \nu$
- $Z \rightarrow \ell \ell$
- Multijet by rebalance & smear
We can say that Higgs decays less than 37% to invisible final states.
Update (2020)
Add 4x more data & improve methods

ATLAS Prelim
\( \sqrt{s} = 13 \text{ TeV}, 139 \text{ fb}^{-1} \)

- Data
- B+S, \( B_{\text{inv}} = 1 \)
- B± syst
- S, \( B_{\text{inv}} = 13\% \)

**Data is consistent with no signal**

- If 100% of the Higgs decayed invisibly
- If 13% of the Higgs decayed invisibly

**Systematic errors**
- simulation samples 8%
- multijet estimation 7%
- jet energy 6%

**Statistical errors**
- large for sensitive bins
- need more data 17%

Two-jet invariant mass

We can say that Higgs decays less than 13% to invisible final states
Interpretations
ATLAS result interpreted in #1, #2

ATLAS result

Interpretation

Quantity

Connection to astrophysics & BSM sector
Interpretation #1
ATLAS result + Higgs portal

scalar: \( \Gamma_{\text{inv}} \cdot (m_\chi)^{-2} \sim \sigma_{\text{WIMP}} \)

fermion: \( \Gamma_{\text{inv}} \cdot \text{const} \sim \sigma_{\text{WIMP}} \)

DM interpretation is complementary to direct detection
Interpretation #2
ATLAS result + "No model"

\[ \sigma_H \cdot \frac{\Gamma_{\text{inv}}}{\Gamma_H} \]

\[ \frac{\varepsilon_{125}}{\varepsilon_{\text{scalar}}} \]

\[ = \sigma_{\text{scalar}} \cdot \frac{\Gamma_{\text{inv}}}{\Gamma_H} \]

\[ = \sigma_{\text{scalar}} \cdot B_{\text{inv}} \]

\[ p \rightarrow H \rightarrow \chi^0, \chi^0 \]

ATLAS Preliminary
\( \sqrt{s} = 13 \text{ TeV}, 139 \text{ fb}^{-1} \)
Limits at 95% CL

TeV scale particle limit at \( \sim \frac{1}{4} \text{ pb} \)
Add a photon to it
Interpretation

$H_{125}$ to $\gamma\gamma_{\text{dark}}$


Data is consistent with no signal
- need more data

If Higgs decayed to dark photons

Observed limit $B_{\text{dark}}$ of 0.014
Expected limit $B_{\text{dark}}$ of 0.017 ± 0.006
Interpretation
Scalar particle to $\gamma \gamma_{\text{dark}}$

ATLAS Preliminary
$\sqrt{s} = 13$ TeV, 139 fb$^{-1}$
Limits at 95% CL
VBF SM-like Higgs couplings

Exclude at few percent up to a few TeV
Trigger

$E_{T\text{miss}}$, VBF
ML on FPGA
**$E_T^{\text{miss}}$ trigger rate vs. $\langle \mu \rangle$**

Rate blows-up with $\langle \mu \rangle$, so very large at the beginning of runs

**Physics**
- Signal: Higgs $p_T$ gives rise to $E_T^{\text{miss}}$
- Background: Combinatorics esp. vs. $\langle \mu \rangle$

**Challenge**
- Problem: $E_T^{\text{miss}}$ trigger is bandwidth limited
- Non-solution: Can't increase threshold to reduce the rate beyond $\sim 150$ GeV bec. signal peaks at low values
- Solution: Be smarter about reducing background while maintaining signal (sounds like physics analysis!)

NB. $150 \rightarrow 180$ GeV reduces signal by $\sim 30\%$
Lots of trouble with pileup driving up MET rates

- Conceptually, linear rate v. $<\mu>$ means “no pileup dependence,” see left
- Rates show non-linear $<\mu>$ dependence, see right plot

Despite periods with very high L1 rates, we kept XE50

- The total rate did peak to almost 8 kHz a few times, but this was only at the start of
At level-1: Smarter noise cuts

Frequent adjustments to noise cuts control the rate

• Adjusted as the pileup increases or the filling scheme changed (right)
• The noise cuts were adjusted three times in 2017 and once in 2018
• Plot on the left shows the impact of the first change (sorry it’s confusing…)

• Documented by L1Calo (link, ATR-17844), but lots of missing changes…

**Noise cuts to reduce L1 rate**

**0.5-1 GeV threshold adjustment in FCAL**

**ATLAS Preliminary**

Fills 5880, 5883

\[ N_{\text{bunches}} = 2554 \]

L1_XE50

- loose FCAL noise threshold
- tight FCAL noise threshold

**Improvement**

**Signal efficiency similar** (before / after noise update with lower rates)
**At HLT: Smarter algorithm (2016)**

*mht110* (default for post-CHEP 2016)

- Rate for $\mu > 45-50$ too high, see left plot

We found backup: *mht110 + cell70*, but kept *mht110*

- The performance is much better when mht is combined with cell MET
- Efficiency is better compared to *mht130*, see right plot

\[
\text{HLT cross-section} = \frac{\text{rate}}{\text{lumi}}
\]

[Efficiency relative to mht110]


Try combo to reduce rate in 2016
At HLT: Smarter (2017)

\( \chi^2 \) based "pileup fit" algorithm

Algorithm

- Divide \( \eta-\phi \) space in \( \sim 0.4^2 \) grid
- Assume uniform underlying pileup energy in \( \eta-\phi \), float magnitude given momentum conservation in \( xy \)

Result

- Trigger rate drastically reduced
- Signal efficiency is similar

Algorithm & threshold evolution

- Rapid development

<table>
<thead>
<tr>
<th>Year</th>
<th>Trigger name</th>
<th>HLT algorithm</th>
<th>L1 threshold [GeV]</th>
<th>HLT threshold [GeV]</th>
<th>( \int L , \mathrm{d}t ) [fb^{-1}]</th>
</tr>
</thead>
<tbody>
<tr>
<td>2015</td>
<td>HLT_xe70_mht_L1XE50</td>
<td>mht</td>
<td>50</td>
<td>70</td>
<td>3.5</td>
</tr>
<tr>
<td>2016</td>
<td>HLT_xe90_mht_L1XE50</td>
<td>mht</td>
<td>50</td>
<td>90</td>
<td>12.7</td>
</tr>
<tr>
<td>2016</td>
<td>HLT_xe110_mht_L1XE50</td>
<td>mht</td>
<td>50</td>
<td>110</td>
<td>30.0</td>
</tr>
<tr>
<td>2017</td>
<td>HLT_xe90_pufit_L1XE50</td>
<td>pufit,cell</td>
<td>50</td>
<td>90,50</td>
<td>21.8</td>
</tr>
</tbody>
</table>

Rapid algorithm R&D to retain \( \text{E}_T^{\text{miss}} \) threshold
At HLT: Combo (2018)
Use combinations of algorithms

Algorithm

• $\chi^2$ based algorithm from prev. slide
• Cell-based algorithm using ~200k LAr cells
• Use both algorithms!

Result

• Trigger rate drastically reduced
• Signal efficiency is similar

Algorithm & threshold evolution

• Rapid development

### Table: Trigger names and thresholds

<table>
<thead>
<tr>
<th>Year</th>
<th>Trigger name</th>
<th>HLT algorithm</th>
<th>L1 threshold</th>
<th>HLT threshold</th>
<th>$\int L , dt$ [fb$^{-1}$]</th>
</tr>
</thead>
<tbody>
<tr>
<td>2015</td>
<td>HLT_xe70_mht_L1XE50</td>
<td>mht</td>
<td>50</td>
<td>70</td>
<td>3.5</td>
</tr>
<tr>
<td>2016</td>
<td>HLT_xe90_mht_L1XE50</td>
<td>mht</td>
<td>50</td>
<td>90</td>
<td>12.7</td>
</tr>
<tr>
<td>2016</td>
<td>HLT_xe110_mht_L1XE50</td>
<td>mht</td>
<td>50</td>
<td>110</td>
<td>30.0</td>
</tr>
<tr>
<td>2017</td>
<td>HLT_xe90_pufit_L1XE50(55)</td>
<td>pufit,cell</td>
<td>50</td>
<td>90, 50</td>
<td>21.8</td>
</tr>
<tr>
<td>2017</td>
<td>HLT_xe100_pufit_L1XE50</td>
<td>pufit,cell</td>
<td>50</td>
<td>100, 33</td>
<td>33.0</td>
</tr>
<tr>
<td>2017</td>
<td>HLT_xe110_pufit_L1XE50(55)</td>
<td>pufit,cell</td>
<td>50 (55)</td>
<td>110, 50</td>
<td>47.7</td>
</tr>
<tr>
<td>2018</td>
<td>HLT_xe110_pufit_xe65_L1XE50</td>
<td>pufit,cell</td>
<td>50</td>
<td>110, 65</td>
<td>57.0</td>
</tr>
<tr>
<td>2018</td>
<td>HLT_xe110_pufit_xe70_L1XE50</td>
<td>pufit,cell</td>
<td>50</td>
<td>110, 70</td>
<td>62.6</td>
</tr>
</tbody>
</table>

Rapid algorithm R&D to retain $E_T^{\text{miss}}$ threshold
**$E_T^{\text{miss}}$ trigger**

Summary of the Run 2 history & my outlook on Run 3

**Approach**

- Start data taking
  - higher pileup than before
- Repeat every few months
  - $E_T^{\text{miss}}$ rates are too high
  - threaten to raise thresholds
- Develop & deploy clever solutions

**Obvious question (& answer)**

- Why not pre-develop in advance? Rates are notoriously difficult to simulate

**My view for Run 3**

- Keep a similar theme of innovating on algorithms, combining algorithms as we did in HLT
- May want to use non-$E_T^{\text{miss}}$ triggers for the $VBF + E_T^{\text{miss}}$ (+ soft) analyses

Rapid algorithm R&D to retain $E_T^{\text{miss}}$ threshold
Upgrade of the level-1 architecture
Run-1, 2

- LAr + Tile
  - Analog receivers
    - trigTowers 0.1 x 0.1
  - Digitize
    - autocorr.
    - ped. cor.
    - noise cor.
  - CP
    - e, γ, τ
  - JEP
    - Jets .8 x .8 MET

L1Calo in yellow

- L1Topo
  - Angles
  - Masses
  - L1 accept

Run 3

- LAr
  - Digitize
    - supercells 0.025 x 0.1
  - eFEX
    - e, γ, τ
  - jFEX
    - Small R=0.4 jets, τ, E_T^{miss}
  - gFEX
    - Global quantities, e.g., Large R=1 jets, E_T^{miss}

- Tile
  - Digitize
    - trigTowers 0.1 x 0.1
  - Same as Run-2
    - Jets, MET
    - e, γ, τ

Run 3 architecture
My guess

- We'll start with baseline $E_T^{\text{miss}}$ algorithms in jFEX-gFEX
- We'll take data and probably realize that we need to do better than baseline
- We'll probably improve & add jFEX-gFEX algorithms (like we did before in HLT)
- We'll combine jFEX-gFEX outputs (like we did before in HLT) → use ML?

Nanosecond machine learning event classification with boosted decision trees in FPGA for high energy physics

Department of Physics and Astronomy, University of Pittsburgh, 100 Allen Hall, 3941 O’Hara St., Pittsburgh, PA 15260, U.S.A.
E-mail: tmhong@pitt.edu

Abstract: We present a novel implementation of classification using the machine learning/artificial intelligence method called boosted decision trees (BDT) on field programmable gate arrays (FPGA). The firmware implementation of binary classification requiring 100 training trees with a maximum depth of 4 using four input variables gives a latency value of about 10 ns, independent of the clock speed from 100 to 320 MHz in our setup. The low timing values are achieved by restructuring the BDT layout and reconfiguring its parameters. The FPGA resource utilization is also kept low at a range from 0.01% to 0.2% in our setup. A software package called FWX.MACHINA achieves this implementation. Our intended user is an expert in custom electronics-based trigger systems in high energy physics experiments or anyone that needs decisions at the lowest latency values for real-time event classification. Two problems from high energy physics are considered, in the separation of electrons vs. photons and in the selection of vector boson fusion-produced Higgs bosons vs. the rejection of the multijet processes.

Keywords: Digital electronic circuits; Trigger algorithms; Trigger concepts and systems (hardware and software); Data reduction methods

arXiv ePrint: 2104.03408

© 2021 IOP Publishing Ltd and IOP Publishing

http://fwx.pitt.edu
Machine learning at L1 trigger

**Level** | **Latency**
--- | ---
L1 | $O(1)$ $\mu$s
HLT | $O(1)$ s

**FPGA**
- Custom electronics
- Reduce rate to 100 kHz

**50k CPUs**
- Software
- Reduce rate to 1-10 kHz

- Partial data
- 65 TB/s
  - 99
  - 0
  - yes / no

- 160 GB/s

- 1.5 GB/s
  - yes / no

Isn't there already a package? (yes, now there are 2)

Deep Neural Network

- Popular method for signal vs. background
- But can't be very deep for FPGA, so ~3 "deep" in paper
- \( y = \Theta(M \cdot x + b) \)

  - Fancy activation (limited resource on FPGA)
  - Multiplication

Standard Decision Tree

- Another popular classification
- \( y = \Theta(x < \text{threshold}) \)

  - Boolean
  - No multiplication (bin search problem)

Decision Tree for FPGA

- Smart bit integer precision, bit shifting
- Flattened (also "deep")
- All variables processed in parallel
- One step algorithm, ns fast, tiny footprint
• **Workflow**

• **Optimization**

• **Use bit integer precision**
• Workflow

• Optimization

• Use bit integer precision

• Will discuss next:
  - Tree Flattener
  - Forest Merger
• **Workflow**

• **Optimization**

• **Use bit integer precision**
  E.g., ap_int<8> means the variable is represented by a range from 0 to 255.

  **Transformation**
  \[ c_{\text{int}} = f(c_{\text{float}}) = \frac{c_{\text{float}} - c_{\text{min}}}{c_{\text{max}} - c_{\text{min}}} \cdot (2^N - 1) \]

• **Advantages & subtleties**
  Bit integers represent a wide range without sacrificing float precision
  
  **Pre-evaluate f**
  **Firmware only adds**
  
  \[ f(x_1 + x_2) = f(x_1) + f(x_2) \]

  **Floor operation**
  **Equal up to one bit because of floor**
First step

Depth $i$

Conventional tree structure

Root node

Start

$O_1$

$q_i: x_a \geq c_i$

False

True

Advantages & subtleties

- Cut thresholds & weights determined during training
- Danger of "memorizing" boundaries (overtraining), so must consider a forest
**Decision tree, 2 var example (2)**

- **First step**
  - **Root node**
  - **Depth i**

- **Conventional tree structure**
  - Start
  - False: \( q_i: x_a \geq c_i \) → \( O_1 \)
  - True

- **2d plane: \( x_a \) vs. \( x_b \)**
  - \( x_a \)
  - \( x_b \)
  - \( c_i \)
  - \( c_{ii} \)

- **Full tree**
  - **Root node**
  - **Depth i**
  - **Depth ii**

- **Advantages & subtleties**
  - Deterministic, conventional style
  - Cuts in each axis is not independent of each other, so recursive
Our approach

- Advantages & subtleties
  - Each axis is independent of each other → Bin search problem on a grid
  - Does not scale well for very deep trees (but do you really need it at L1?)

Full tree
Forest of boosted decision trees

- **Advantages & subtleties**
  - Use TMVA software to train the BDT (support for other sw coming)

- Can we pre-merge the trees for firmware? Yes, next slide.

Our approach

**Decision tree** $\tau_\beta$ with boost weight $w_\beta$

- **Decision tree** $\tau_\alpha$ with boost weight $w_\alpha$

---

**1st tree**

- **2nd tree**
Merging of the forest

• **Advantages & subtleties**
  • Merging is pre-processed before implementation in firmware
  • This is using adaptive boosting. Gradient boosting cannot pre-merge, but we have approximations for that method to improve performance.

• **Physics impact of flattening & merging**
  • None, bec. encodes the entirety of conventional approach
  • Firmware is a giant look-up table problem
VBF Higgs vs. Multijet background

- $\sigma_{\text{Higgs}} = 4$ pb, two widely separated high-$p_T$ jets
- $\sigma_{pp} = 80$ mb, dominant process at LHC
- Distributions given on the right

We consider two decays of the Higgs

- $H \rightarrow$ neutrinos, "invisible"
- $H \rightarrow b\bar{b}b\bar{b}$, thru pseudoscalar decays

Strategy

- Train BDT to identify VBF jet pair,
  i.e., train BDT on Multijet vs. VBF $H \rightarrow$ neutrinos
- Apply that BDT to Multijet vs. VBF $H \rightarrow b\bar{b}b\bar{b}$

Why

- If it works for VBF $H \rightarrow b\bar{b}b\bar{b}$, then it can be a trigger for VBF independent of the Higgs decay
- Does it work? Next slide
It works!
- Reminder. Did not train on VBF $H \rightarrow b\bar{b}b\bar{b}$
- Subtlety re: jet selection (see paper)
- Distributions given on the right

Performance comparison
- Try to mimic ATLAS HL-LHC cuts as best we can using Madgraph + Delphes
- Two-fold signal efficiency improvement from ATLAS-inspired $\rightarrow$ fwX results

Details
- We validated our setup to reproduce the signal efficiency in the ATLAS Run-2 paper
- Comparison using bit integers, not floats
Ran two configurations
- Optimized version
- Non-optimized version (for comparison)
- Both using 100 trees, max depth of 4
- Results given on the right

Performance
- 5 clock ticks = 16 ns
- Negligible resource usage

Benchmark using $e^+$ vs. $\gamma$
- In the paper, we also define one set of parameters to scale up one param. at a time
- Uses 4 variables, 8 bits & same as above
- 3 clock ticks = 10 ns
- Negligible resource usage

<table>
<thead>
<tr>
<th></th>
<th>VBF H Optimized</th>
<th>VBF H Non-opt</th>
</tr>
</thead>
<tbody>
<tr>
<td>$N_{\text{var}}$</td>
<td>5</td>
<td>7</td>
</tr>
<tr>
<td>$N_{\text{bit-var}}$</td>
<td>8</td>
<td>12</td>
</tr>
<tr>
<td>$N_{\text{bit-score}}$</td>
<td>16</td>
<td>16</td>
</tr>
<tr>
<td>$N_{\text{bin}}$</td>
<td>40k</td>
<td>1M</td>
</tr>
<tr>
<td>Latency</td>
<td>5 ticks</td>
<td>6 ticks</td>
</tr>
<tr>
<td>LUT</td>
<td>1%</td>
<td>1.5%</td>
</tr>
<tr>
<td>Flip Flops</td>
<td>~0</td>
<td>~0</td>
</tr>
<tr>
<td>BRAM</td>
<td>2%</td>
<td>30%</td>
</tr>
<tr>
<td>DSP</td>
<td>0</td>
<td>~0</td>
</tr>
</tbody>
</table>
**Where can we put fwX?**

**My guess**

- L1Topo to do combo algorithms
- gFEX to develop new algorithms
- jFEX to develop new algorithms
- eFEX to develop new algorithms

**Run 3**

- LAr
  - Digitize supercells $0.025 \times 0.1$
- Tile
  - Digitize trigTowers $0.1 \times 0.1$

**Sources:**
- L1Calo TDR, [https://cds.cern.ch/record/1602235](https://cds.cern.ch/record/1602235)
- This diagram courtesy B. Carlson.
Regression (using BDT)

- Toy problem in 1-d
- Train / test on $f(x) = \sin(x) + \text{Gaussian}(x)$
- For sample of $x$: $y = f(x)$ in 16 bits

<table>
<thead>
<tr>
<th>$N_{depth}$</th>
<th>$N_{tree} = 1$</th>
<th>$N_{tree} = 10$</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td><img src="image" alt="Graph for N_tree = 1, N_depth = 2" /></td>
<td><img src="image" alt="Graph for N_tree = 10, N_depth = 2" /></td>
</tr>
<tr>
<td>4</td>
<td><img src="image" alt="Graph for N_tree = 1, N_depth = 4" /></td>
<td><img src="image" alt="Graph for N_tree = 10, N_depth = 4" /></td>
</tr>
<tr>
<td>(x, y) = (N_{tree}, N_{depth})</td>
<td>N_{tree} = 1</td>
<td>10</td>
</tr>
<tr>
<td>-------------------------------</td>
<td>-------------</td>
<td>----</td>
</tr>
<tr>
<td>2</td>
<td><img src="image1" alt="Graph" /></td>
<td><img src="image2" alt="Graph" /></td>
</tr>
<tr>
<td>4</td>
<td><img src="image6" alt="Graph" /></td>
<td><img src="image7" alt="Graph" /></td>
</tr>
<tr>
<td>8</td>
<td><img src="image11" alt="Graph" /></td>
<td><img src="image12" alt="Graph" /></td>
</tr>
</tbody>
</table>
Summary

Why (10 min)
• Higgs boson
• Dark sector

Method (20 min)
• $H_{125} \rightarrow$ Dark matter pair  $B_{\text{invisible}} < 0.13$
• $H_{125} \rightarrow$ Dark photon + $\gamma$  $B_{\text{dark photon}} < 0.014$

Trigger (30 min)
• $E_{T\text{miss}}$, VBF
• ML on FPGA  http://fwx.pitt.edu

Higgs portal to DM? How to trigger?
Abstract

With more data coming from LHC collisions, detailed measurements of Higgs boson properties allow us to probe whether it communicates with the unknown and/or undiscovered sector beyond the Standard Model. One motivation is weakly interacting dark matter, which are invisible to the detecting apparatus, through a Higgs portal. I will discuss the latest ATLAS results of the search for Higgs bosons decaying to invisible particles. I will also describe the technical challenges of triggering on such events using missing energy from the Higgs boson decay and/or hadronic jets from the Higgs boson production, including the potential use of machine learning methods on FPGA boards in real-time level-1 trigger systems. will discuss how such interactions produce the recently discovered Higgs boson, and how it may serve as a portal to unknown sectors of elementary particles, such as dark matter. I will also describe the technical challenges of saving such minuscule fractions of weak force collisions, including the use of artificial intelligence in real-time trigger systems.
Distribution of energy
Simulation of the polar angle for one collision

Ellis, Huston, Hatakeyama, Loch, Tönnesmann, Prog. in Part. & Nucl. Phys. 60 (2008) 484

Large $\Delta\eta$ between the scattered quark jets
Detector signature

- **ATLAS geometry**
  - $\eta$ along the beam direction
  - $\phi$ azimuthal angle

- **VBF jet pair**
  - High $p_T$
  - Wide gap in $\eta$
  - Not back-to-back in $\phi$
  - Large $m_{\text{jet-jet}}$ \(2\) TeV $\rightarrow$
  - Low hadronic activity in between

- **$E_T^{\text{miss}}$**
  - $p_T$ imbalance \(840\) GeV $\rightarrow$

- **For $+\gamma$**
  - High-$p_T$ photon \(540\) GeV $\rightarrow$
  - $m_T(E_T^{\text{miss}}, \gamma)$ \(1.1\) TeV $\rightarrow$

- Its $\eta$ in between jets

Signal models

• H portal to $\chi$
  - VBF $H_{125}$ w/ POWHEG NLO
  - VBF $H_{125} + \gamma_{ISR}$ w/ MG5_aMC@NLO
  - S-to-B is higher with $m_{jj}$, $E_{T}^{\text{miss}}$, see →

• H portal to $\gamma_d$
  - VBF $H_{125} \rightarrow \gamma\gamma_{\text{dark}}$ w/ POWHEG v2
  - $m_T$($E_{T}^{\text{miss}}$, $\gamma$) as proxy for $m_H$, see →

http://cern.ch/Atlas/GROUPS/PHYSICS/PAPERS/EXOT-2016-37/fig_05.pdf


Transverse mass of $E_{T}^{\text{miss}}$ and $\gamma$

ATLAS Simulation Preliminary
\( \sqrt{s} = 13 \text{ TeV}, 139 \text{ fb}^{-1} \)

Reco

Truth

0 200 400 600 800 1000

m_T [GeV]

0.05 0.1 0.15 0.2

Unit Normalized

60 GeV 125 GeV 500 GeV 1000 GeV

• Weak boson bkg'd
  • $Z \rightarrow \nu \nu$  No leptons
  • $W \rightarrow \ell \nu$  Loses a lepton

• Signal Region
  • $E_T^{\text{miss}}$ trigger, $> 150$ GeV
  • "Centrality" of $\gamma$, 3rd jet
  • For $+\gamma_{\text{ISR}}$, $15 < p_T^\gamma < 110$ GeV
  • For $+\gamma_{\text{dark}}$, $\max(110,0.7 m_T)$

• Control Region
  • For $W \rightarrow \ell \nu$, Require a lepton
  • Lepton trigger, $> 30$ GeV
  • Reverse $\gamma$ centrality cut

57

Uncertainties, for $\gamma\gamma$

- **Statistical**
  - $\sqrt{N}$
  - MC

- **Theoretical**
  - $W\gamma$, $Z\gamma$ theory

- **Experimental**
  - JES, JER

<table>
<thead>
<tr>
<th></th>
<th>Data stats.</th>
<th>$V\gamma$+ jets theory</th>
<th>MC stats.</th>
<th>Jet Scale and Resolution</th>
</tr>
</thead>
<tbody>
<tr>
<td>Photon</td>
<td>0.032</td>
<td>0.032</td>
<td>0.045</td>
<td>0.045</td>
</tr>
<tr>
<td>$e \rightarrow \gamma$, jet$\rightarrow e,\gamma$ Bkg.</td>
<td>0.026</td>
<td>0.026</td>
<td>0.045</td>
<td>0.045</td>
</tr>
<tr>
<td>Pileup</td>
<td>0.025</td>
<td>0.025</td>
<td>0.045</td>
<td>0.045</td>
</tr>
<tr>
<td>$W\gamma$+ jets/$Z\gamma$+ jets Norm.</td>
<td>0.021</td>
<td>0.021</td>
<td>0.045</td>
<td>0.045</td>
</tr>
<tr>
<td>$E_T^{miss}$</td>
<td>0.012</td>
<td>0.012</td>
<td>0.045</td>
<td>0.045</td>
</tr>
<tr>
<td>Signal theory</td>
<td>0.004</td>
<td>0.004</td>
<td>0.045</td>
<td>0.045</td>
</tr>
<tr>
<td>Lepton</td>
<td>0.002</td>
<td>0.002</td>
<td>0.045</td>
<td>0.045</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>0.148</strong></td>
<td><strong>0.148</strong></td>
<td><strong>0.106</strong></td>
<td><strong>0.106</strong></td>
</tr>
</tbody>
</table>

$1\sigma$ Uncertainty on $B_{inv}$ on $B(H \rightarrow \gamma\gamma_d)$

Evaluated by fixing parameters to their best-fit values and quadratically subtracting from the total nominal systematic uncertainty.

---

Table 1: Summary of generators used for simulation. The details and the corresponding references are provided in the body of the text. The $V$ in $V$+jets represents either a $W$ or a $Z$ boson.

<table>
<thead>
<tr>
<th>Process</th>
<th>Generator</th>
<th>ME Order</th>
<th>PDF</th>
<th>Parton Shower</th>
<th>Tune</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Signal Samples</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ggF Higgs</td>
<td>Powheg v2 NNLOPS</td>
<td>NNLO</td>
<td>PDF4LHC15</td>
<td>Pythia8.230</td>
<td>AZNLO</td>
</tr>
<tr>
<td>VBF Higgs + $\gamma$</td>
<td>MadGraph5_aMC@NLO 2.6.2</td>
<td>NLO</td>
<td>PDF4LHC15</td>
<td>Herwig 7.1.3p1</td>
<td>A14</td>
</tr>
<tr>
<td>ggF Higgs $\rightarrow$ $\gamma\gamma_d$</td>
<td>Powheg v2 NNLOPS</td>
<td>NNLO</td>
<td>PDF4LHC15</td>
<td>Pythia8.244p3</td>
<td>AZNLO</td>
</tr>
<tr>
<td>VBF Higgs $\rightarrow$ $\gamma\gamma_d$</td>
<td>Powheg v2</td>
<td>NLO</td>
<td>CTEQ6L1</td>
<td>Pythia8.244p3</td>
<td>AZNLO</td>
</tr>
<tr>
<td><strong>Background Samples</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Strong $V\gamma$+jets</td>
<td>Sherpa v2.2.8</td>
<td>NLO</td>
<td>NNPDF3.0nnlo</td>
<td>Sherpa MEPS@NLO</td>
<td>Sherpa</td>
</tr>
<tr>
<td>EW $V\gamma$+jets</td>
<td>MadGraph5_aMC@NLO 2.6.5</td>
<td>LO</td>
<td>NNPDF3.11o</td>
<td>Pythia8.240</td>
<td>A14</td>
</tr>
<tr>
<td>EW $VV$+jets</td>
<td>Sherpa v2.2.1 or Sherpa v2.2.2</td>
<td>LO</td>
<td>NNPDF3.0nnlo</td>
<td>Sherpa MEPS@LO</td>
<td>Sherpa</td>
</tr>
<tr>
<td>$VV$+jets</td>
<td>Sherpa v2.2.1 or Sherpa v2.2.2</td>
<td>NLO (up to 1-jet), NNPDF3.0nnlo</td>
<td>Sherpa MEPS@NLO</td>
<td>Sherpa</td>
<td></td>
</tr>
<tr>
<td>EW $V$+jets</td>
<td>Herwig 7.1.3 or Herwig 7.2.0</td>
<td>NLO</td>
<td>MMHT2014nlo68cl</td>
<td>Herwig 7.1.3</td>
<td>Herwig 7</td>
</tr>
<tr>
<td>Strong $W(\rightarrow \mu\nu)$+jets/ $W(\rightarrow \tau\nu)$+jets</td>
<td>Sherpa v2.2.8</td>
<td>NLO (up to 2-jets), NNPDF3.0nnlo</td>
<td>Sherpa MEPS@NLO</td>
<td>Sherpa</td>
<td></td>
</tr>
<tr>
<td>$t\bar{t}\gamma$</td>
<td>MadGraph5_aMC@NLO 2.2.3</td>
<td>NLO</td>
<td>NNPDF3.2.31o</td>
<td>Pythia8.186</td>
<td>A14</td>
</tr>
<tr>
<td>$t\bar{t}$</td>
<td>PowhegBox v2</td>
<td>NLO</td>
<td>NNPDF3.0nnlo</td>
<td>Pythia8.230</td>
<td>A14</td>
</tr>
<tr>
<td>$\gamma$ + jet</td>
<td>Sherpa v2.2.2</td>
<td>NLO (up to 2-jets), NNPDF3.0nnlo</td>
<td>Sherpa MEPS@NLO</td>
<td>Sherpa</td>
<td></td>
</tr>
<tr>
<td><strong>Systematic Samples</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$V\gamma$+jets $\alpha^4$</td>
<td>MadGraph5_aMC@NLO 2.6.2</td>
<td>LO</td>
<td>NNPDF3.11o</td>
<td>Pythia8.240</td>
<td>AZNLO</td>
</tr>
</tbody>
</table>

Signal region for $+\gamma$

Table 3: Summary of the requirements defining the different regions considered in this analysis. Where present, the values in squared brackets are referring to the regions defined in the search for $H \rightarrow \gamma\gamma_d$ signal. The leading and subleading jets must satisfying the fJVT requirements mentioned in Sec. 5. In the SR and $Z_{\text{Rev. Cen.}}^{\gamma}$ CR definitions $E_{T}^{\text{miss,lep-rm}} \equiv E_{T}^{\text{miss}}$ since no lepton is present.

<table>
<thead>
<tr>
<th>Variable</th>
<th>SR</th>
<th>$W_{\mu\nu}\gamma$ CR</th>
<th>$W_{e\nu}\gamma$ CR</th>
<th>$Z_{\text{Rev. Cen.}}^{\gamma}$ CR</th>
<th>Fake-$e$ CR</th>
</tr>
</thead>
<tbody>
<tr>
<td>$p_T (j_1)$ [GeV]</td>
<td>$&gt; 60$</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$p_T (j_2)$ [GeV]</td>
<td>$&gt; 50$</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$N_{\text{jet}}$</td>
<td>$2,3$</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$N_{\text{b-jet}}$</td>
<td>$&lt; 2$</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\Delta\phi_{jj}$</td>
<td>$&lt; 2.5$</td>
<td>$[2.0]$</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$</td>
<td>\Delta\eta_{jj}</td>
<td>$</td>
<td>$&gt; 3.0$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\eta(j_1) \times \eta(j_2)$</td>
<td>$&lt; 0$</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$C_3$</td>
<td>$&lt; 0.7$</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$m_{jj}$ [TeV]</td>
<td>$&gt; 0.25$</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$E_{T}^{\text{miss}}$ [GeV]</td>
<td>$&gt; 150$</td>
<td>$–$</td>
<td>$&gt; 80$</td>
<td>$&gt; 150$</td>
<td>$&lt; 80$</td>
</tr>
<tr>
<td>$E_{T}^{\text{miss,lep-rm}}$ [GeV]</td>
<td>$–$</td>
<td>$&gt; 150$</td>
<td>$&gt; 150$</td>
<td>$–$</td>
<td>$&gt; 150$</td>
</tr>
<tr>
<td>$E_{T}^{\text{jets,no-jvt}}$ [GeV]</td>
<td>$&gt; 130$</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\Delta\phi(j_i, E_{T}^{\text{miss,lep-rm}})$</td>
<td>$&gt; 1.0$</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$N_{\gamma}$</td>
<td>$1$</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$p_T (\gamma)$ [GeV]</td>
<td>$&gt; 0.4$</td>
<td>$&gt; 0.4$</td>
<td>$&gt; 0.4$</td>
<td>$&lt; 0.4$</td>
<td>$&gt; 0.4$</td>
</tr>
<tr>
<td>$C_{\gamma}$</td>
<td>$&gt; 0.4$</td>
<td>$&gt; 0.4$</td>
<td>$&gt; 0.4$</td>
<td>$&lt; 0.4$</td>
<td>$&gt; 0.4$</td>
</tr>
<tr>
<td>$\Delta\phi(\gamma, E_{T}^{\text{miss,lep-rm}})$</td>
<td>$&gt; 1.8$ [–]</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$N_{\ell}$</td>
<td>$0$</td>
<td>$1 \mu$</td>
<td>$1 e$</td>
<td>$0$</td>
<td>$1 e$</td>
</tr>
<tr>
<td>$p_T (\ell)$ [GeV]</td>
<td>$&gt; 30$</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

$\eta_{\gamma}, \eta_{j2}$

Centrality $C_{\gamma}$ [102] is defined as

$$C_{\gamma} = \exp\left(-\frac{4}{(\eta_{1} - \eta_{2})^2 (\eta_{\gamma} - \frac{\eta_{1} + \eta_{2}}{2})^2}\right). \quad (1)$$