DARK MATTER SEARCH ON THE MONO-HIGGS (H→bb) IN THE ATLAS EXPERIMENT

DILIA MARIA PORTILLO QUINTERO

30th October 2018 - Puzzle of Dark Matter
DESY Hamburg
HOW TO IMPROVE MONO-H ANALYSIS?

Main ingredients:
- Jets
- b-tagging
- MET

Commissioning and exercising **NEW reconstructions techniques**

Object based MET Significance

Object-based MET Significance helps to reduce the multi-jet background and/or improve its estimation

Variable Radius track jets

VR track jets helps to increase the sensitivity in boosted topologies with b-tagging.
**MISSING TRANSVERSE ENERGY RECONSTRUCTION**

- **Missing Transverse Energy (MET)**

\[
\sum_i \vec{p}_i = 0 \quad \text{When something is missing}
\]

\[
\sum_{\text{observable}} \vec{p}_i + \vec{E}_T^{\text{miss}} = 0 \quad \vec{E}_T^{\text{miss}} = -\sum_{\text{observable}} \vec{p}_i
\]

Two effects could imply an imbalance in the total transverse momentum (\(\text{MET} \neq 0\))

- **Non-interacting particles** (Real MET): SM particles (neutrinos), new physics (Dark Matter, SUSY, etc.)
- **Fake detection** (Fake MET): Objects misreconstruction, detector effects (dead regions).

Reconstructed Met = Calorimeter Objects + Muons + Soft Terms

- Electrons, Photons, Taus, Jets
- Hard Objects

Every signal which can’t be clearly identified/calibrated.
Significance of an observable

\[
\text{Signif}(X) = \frac{X}{\sigma_X}
\]

**MET significance** helps separate events with **real MET** versus events with **fake MET**

**Event-based MET Significance**

\[
S = \frac{E_{\text{miss}}}{\sqrt{\sum E_T}}
\]

- Proxi for the MET resolution
- Event based quantity
- Does not take into account directional correlations

**Object-based MET Significance Definition**

- Based on the **expected resolutions** for all objects that enter the MET reconstruction
- **Event by event** calculated

\[
S^2 = 2 \ln \left( \frac{\mathcal{L}(E_{T \text{miss}} | E_{T \text{miss}})}{\mathcal{L}(E_{T \text{miss}} | 0)} \right) = (E_{T \text{miss}})^T \left( \sum_i V_i \right)^{-1} (E_{T \text{miss}})
\]

This novel definition depends on the multiplicities, types, and kinematics of the objects measured in each event

- Covariance Matrix for each object
  - Hard objects
  - Soft term
The multijet background is poorly described by MC, requiring a data-driven estimate.

Multijet background introduces fake MET mainly due to mis-measured jet momenta.

Met significance can identify and separate multijet background with respect to EW backgrounds and DM signals.

Object-based MET significance in mono-h(bb) search

New requirement in resolved selection to suppress multijet background (poorly modelled by the dijet MC) at low values of the distribution.
Object-based MET significance separation power

- Signal: Z’-2HDM, \((m_{Z’}, m_A) = (400, 300)\) GeV
- Background: Dijet MC

Expected signal significance with a MET significance/without MET significance requirement

Expected signal significance not compromised with Met significance > 16 requirement

\[
\text{Signal: } Z' - 2\text{HDM, (}m_{Z'}, m_A) = (400, 300)\text{ GeV}
\]

\[
\text{Background: Dijet MC}
\]

With an object-based MET significance>16 requirement:
- More than 95% of dijet can be rejected while retaining a signal efficiency ~90%
- Signal significance gain between 6%-9%
**RESIDUAL MULTIJET**

Estimate residual multijet event yield after object-based MET significance requirement with ABCD method:

- Define regions A,B,C by cut values on \( \min \Delta \phi (j_1, j_2, j_3, \text{MET}) \) and MET Significance

<table>
<thead>
<tr>
<th>Region</th>
<th>Data</th>
<th>Background</th>
<th>ABCD prediction</th>
</tr>
</thead>
<tbody>
<tr>
<td>150 GeV &lt; ( E_T^{miss} ) &lt; 200 GeV</td>
<td>7698 ± 88</td>
<td>6397 ± 45</td>
<td>20 ± 3.3</td>
</tr>
<tr>
<td>200 GeV &lt; ( E_T^{miss} ) &lt; 350 GeV</td>
<td>5429 ± 74</td>
<td>4731 ± 36</td>
<td>8 ± 2.9</td>
</tr>
<tr>
<td>350 GeV &lt; ( E_T^{miss} ) &lt; 500 GeV</td>
<td>484 ± 22</td>
<td>434 ± 9</td>
<td>0.1 ± 0.76</td>
</tr>
</tbody>
</table>

Multijet yield prediction for signal region (SR):

\[
N_A = \frac{N_B N_C}{N_D}
\]

Negligible since smaller than data statistical uncertainty.
Merging of Fixed Radius (FR) track jets causes loss of acceptance $\times$ efficiency for signals with highly boosted topologies.

The reconstruction of sub-jets used for b-tagging in the merged regime improves using the VR track jets, resulting in a **higher b-tagging efficiency**.

**VR track-jets**

Anti-kt algorithm with variable radius parameter

$$d_{ij} = \min \left( \frac{\Delta R_{ij}^2}{R_{\text{eff}}(p_T(i))^2} \right)$$

$$R_{\text{eff}}(p_T(i)) = \frac{\rho}{p_T(i)}$$

arXiv:0903.0392  \hspace{1cm} R = 0.02-0.4,  \rho = 30 \text{ GeV}
A novel definition of MET significance has been developed. It depends on the multiplicities, types, and kinematics of the objects measured in each event.

The new definition has a clear improvement in the separation power between samples with genuine MET and samples with no genuine MET.

Improved search for dark matter with 79.8 fb\(^{-1}\), commissioning and validating novel reconstruction techniques: VR track jets and object-based MET significance.

Object-based MET significance successfully reject multijet background maintaining signal significance.

3× improvement driven by VR tracks-jets in boosted region!

[ATLAS-CONF-2018-038]

[ATLAS-CONF-2018-039]
BACKUP
Jets represent hadronic showers in a detector.

ATLAS primarily uses the anti-kt algorithm, with topo-cluster inputs: **Calorimeter jets**
- Topo-clusters: topological groups of noise-suppressed calorimeter cells

Depending on physics intent, different types of jets are useful:
- **Small-R jets: R = 0.4**
  - For boosted topologies:
    - **Large-R jets: R = 1.0**

**B-jet identification: “b-tagging”**

MVA algorithm. Uses b-hadron properties:
- Large lifetime
- High decay multiplicity
- Often decays semi-leptonically ~42%
Hierarchical clustering algorithm that has a resolution parameter, R, which controls the extension of the jet

\[ d_{ij} = \min(k_{ti}^{2p}, k_{tj}^{2p}) \frac{\Delta_{ij}^2}{R^2} \]

\[ d_{iB} = k_{ti}^{2p} \]

\[ \Delta_{ij}^2 = (y_i - y_j)^2 + (\phi_i - \phi_j)^2 \]

d\(_{ij}\): distances between entities (particles, pseudojets) i and j
d\(_{iB}\): distance between entity i and the beam (B)

Algorithm:
- Calculate the distances between the particles: \( d_{ij} \)
- Calculate the beam distances: \( d_{iB} \)
- Combine particles with smallest distance \( d_{ij} \) or, if \( d_{iB} \) is smallest, call it a jet
- Find again smallest distance and repeat procedure until no particles are left

Anti-kt: If a hard particle has no hard neighbours within a distance 2R, then it will simply accumulate all the soft particles within a circle of radius R, resulting in a perfectly conical jet.
Separate b-jets from light (u,d,s,g) and charm jets using specific b-hadron properties:

- Mass of b-hadrons (5 GeV)
- Large lifetime (~1.5 ps) → Secondary Vertex and tracks with large IP.
- In ~42% of the cases the b-hadron decays semi-leptonically, in ~11% directly (b → ℓ) and in ~10% indirectly (b → c → ℓ). → search for “soft” muons in the SV.
combined into a single high-level Tagger: MV2C10

The algorithm trains the MV2c10 with b-jets as signal, while the background sample is composed of 7% (93%) c- (light-flavour) jets in $t\bar{t}$ events.

![Graph showing ATLAS Simulation Preliminary data for $\sqrt{s} = 13$ TeV, $t\bar{t}$]

<table>
<thead>
<tr>
<th>BDT Cut Value</th>
<th>$b$-jet Efficiency [%]</th>
<th>$c$-jet Rejection</th>
<th>Light-jet Rejection</th>
<th>$\tau$ Rejection</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.9349</td>
<td>60</td>
<td>34</td>
<td>1538</td>
<td>184</td>
</tr>
<tr>
<td>0.8244</td>
<td>70</td>
<td>12</td>
<td>381</td>
<td>55</td>
</tr>
<tr>
<td>0.6459</td>
<td>77</td>
<td>6</td>
<td>134</td>
<td>22</td>
</tr>
<tr>
<td>0.1758</td>
<td>85</td>
<td>3.1</td>
<td>33</td>
<td>8.2</td>
</tr>
</tbody>
</table>
MONO-H SIMPLIFIED MODELS

Vector+Pseudoscalar (Z'-2HDM)

**New massive vector boson**

**Pseudoscalar from a 2HDM Type 2**

**Different pseudoscalar masses** $m_A$

$\begin{align*}
\text{Different} & \quad \text{masses} \quad m_A \\
\text{events} & \quad m_A = 800 \text{ GeV} \\
\text{events} & \quad m_A = 300 \text{ GeV}
\end{align*}$

**Fixed Parameters**

- $m_\chi = 100 \text{ GeV}$
- $\tan \beta = 1.0$
- $g_{Z0} = 0.8$
- $m_H = m_{H^\pm} = 300 \text{ GeV}$
- $\text{BR}(A \to \chi\chi) = 100\%$

**Different dark matter masses** $m_{DM}$

**Z'-2HDM, $m_Z = 1 \text{ TeV}, m_{A^0} = 300 \text{ GeV}**

$\begin{align*}
\text{events/20 GeV} & \\
\text{events/20 GeV} & \text{arXiv: 1507.00966}
\end{align*}$
Search for DM produced in association with Higgs→bb

Event selection overview:
- Trigger on high MET
- No leptons (e/µ/τ veto)
- Identify $h→bb$ decay

Look for excess in reconstructed Higgs mass $m_{jj}/m_J$

Resolved Regime
- Higgs is reconstructed with a pair of small radius jets (j)

Merged Regime
- For boosted events, the Higgs is reconstructed with a large radius jet (J) with substructure

$150 < \text{MET} < 500 \text{ GeV}$

$\text{MET} > 500 \text{ GeV}$
Improve the gain in selection efficiency for large $m_{Z'}$.

**Good complementarity of two regimes**
DISTRIBUTIONS AFTER FIT

No excess in data observed
⇒ constrain model parameters with exclusion limit contour
Model parameters excluded at 95%CL:

\[ m_{Z'} \leq 2.8 \text{ TeV} \] and \[ m_A \leq 600 \text{ GeV} \]

The improvement from using VR track jets

\[ 36 \text{fb}^{-1} \text{ using FR] } \]

The improvement from using VR track jets

\[ 3\times \text{ improvement driven by VR tracks-jets in boosted region!} \]
MISSING TRANSVERSE ENERGY PERFORMANCE

In order to study the performance of the MET, we want to consider a clean process without genuine MET: \( Z \rightarrow \text{leptons} + \text{jet} \).

The bulk of the distribution mostly provided by fake MET reconstruction.

MET Resolution

\[ \text{MET Resolution} \]
BEYOND THESE APPROXIMATIONS: OBJECT-BASED MET SIGNIFICANCE

Object-based MET significance definition
◇ Based on the expected resolutions for all objects that enter the MET reconstruction
◇ Event by event calculated

How likely is the measured MET to be genuine MET or a resolution effect?
This can be evaluated with the log-likelihood ratio:

\[ S^2 = 2 \ln \left( \frac{\max_{p_T^{inv} \neq 0} \mathcal{L}(E_T^{miss} | p_T^{inv})}{\max_{p_T^{inv} = 0} \mathcal{L}(E_T^{miss} | p_T^{inv})} \right) \]

If we assume that …
- The sum of all the truth transverse momentum is equal to zero
- Gaussian object resolutions

Met Significance corresponds to a \( \chi^2 \) variable:

\[ S^2 = \left( E_T^{miss} \right)^T \left( \sum_i V_i^i \right)^{-1} E_T^{miss} \]

In a coordinate system longitudinal and transverse to the total MET axis:

\[ \sigma_T \]
\[ \sigma_L \]
\[ \sigma_y \]
\[ \sum_i V_i^i = \begin{pmatrix} \sigma_L^2 & \rho_L T \sigma_L \sigma_T \\ \rho_L T \sigma_L \sigma_T & \sigma_T^2 \end{pmatrix} \]
OBJECT-BASED MET SIGNIFICANCE DEFINITION

\[ S = \frac{|E_T^{\text{miss}}|}{\sqrt{\sigma_L^2 (1 - \rho_{LT}^2)}} \]

Total Variance Longitudinal
\[ \sigma_L^2 = \sigma_{L_{\text{hard}}}^2 + \sigma_{L_{\text{soft}}}^2 \]

Transverse momentum resolution in the transverse plane for the hard objects
\[ \mathbf{v}_i = \begin{pmatrix} \sigma_{p_T i}^2 & 0 \\ 0 & p_T i^2 \sigma_{\phi i}^2 \end{pmatrix} \]

Resolution parallel to the direction of the object
Resolution perpendicular to the direction of the object

Resolutions
- Hard object variances
- Soft Terms variance

**ATLAS Simulation Preliminary**

<table>
<thead>
<tr>
<th>Object</th>
<th>Kinematic</th>
<th>Relative resolution</th>
<th>(\rho_T) [GeV]</th>
<th>(\eta)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Electrons</td>
<td>(p_T = 100) GeV, (\eta = 0)</td>
<td>1.7%</td>
<td>0.4%</td>
<td>0.0</td>
</tr>
<tr>
<td>Photons</td>
<td>(p_T = 100) GeV, (\eta = 0)</td>
<td>1.9%</td>
<td>0.4%</td>
<td>0.0</td>
</tr>
<tr>
<td>Hadronic (\tau)</td>
<td>(p_T = 100) GeV, (\eta = 0)</td>
<td>5.5% – 6.7%</td>
<td>1%</td>
<td>0.0</td>
</tr>
<tr>
<td>Jets</td>
<td>(p_T = 20) GeV, (\eta = 0)</td>
<td>22%</td>
<td>4.6%–7.1%</td>
<td>0.0</td>
</tr>
<tr>
<td>Muons</td>
<td>(p_T = 100) GeV, (\eta = 0)</td>
<td>7%</td>
<td>1.1%–1.6%</td>
<td>0.0</td>
</tr>
<tr>
<td>Track Soft Term</td>
<td>(p_T = 100) GeV, (\eta = 0)</td>
<td>2–4%</td>
<td>0.1%</td>
<td>0.0</td>
</tr>
</tbody>
</table>

Total Variance Longitudinal

24
OBJECT-BASED MET SIGNIFICANCE DEFINITION

\[
S = \frac{|E_T^{\text{miss}}|}{\sqrt{\sigma_L^2 (1 - \rho_{LT}^2)}}
\]

**Resolutions**
- Hard object variances
- Soft Terms variance

**Total Variance Longitudinal**
\[
\sigma_L^2 = \sigma_{hard}^2 + \sigma_{soft}^2
\]

An additional **covariance matrix** associated to the soft term is also considered:

\[
V_{soft} = \begin{pmatrix}
\sigma_{soft}^2 & 0 \\
0 & \sigma_{soft}^2
\end{pmatrix}
\]

**Constant soft Term resolution**

Improving object-based MET significance

The neutrals not associated to any hard objects are not in the Soft Term. As a result we expect to have a **bias**.

The **8.9 GeV resolution** mixes the effect of the intrinsic Soft Term **resolution** and the presence of the **bias**.
A jet from pile-up is unbalancing the event, and creating a fake MET. Effect on the MET is equal to the jet pT.

Estimation of the effect of these jets is done from MC simulation, and relative pT resolution parametrised to be equal to the fraction of jets from pile-up in a certain region of pT, $|\eta|$, JVT score.
Comparison of the separation power between
- Background: $Z \rightarrow ee + \text{jet}$
- Signal: $ZZ \rightarrow ee\nu\nu + \text{jet}$
PERFORMANCE STUDY

- Comparison of the separation power between
  - Background: $Z \rightarrow ee + \text{jets}$ → No genuine MET
  - Signal: $ZZ \rightarrow ee\nu\nu + \text{jets}$ → Neutrinos!

![Object-based MET Significance](image1)

![Event-based MET Significance](image2)
Comparison of the separation power between
- **Background:** $Z \rightarrow ee + \text{jets}$  ➔ No real MET
- **Signal:** $ZZ \rightarrow eevv + \text{jets}$  ➔ Neutrinos!

**MET > 100 GeV**
- **Object-based MET significance**

**1 Jet, MET > 50 GeV**
- **Event-based MET significance**

Significant improvements!
**MONO-H CONTROL REGIONS**

**Backgrounds**

- **Irreducible** $Z_h \rightarrow \nu \nu + bb$
- **Dominant** $Z \rightarrow \nu \nu + \text{Jets}$
- $W \rightarrow l(e/\mu)\nu$
- $t\bar{t}$ +production

**Signal and Control Regions**

- **Signal Region: 0-lepton (SR)**
- **Control Regions:**
  - 1-lepton ($1\mu$-CR)
  - 2-lepton ($2\ell$-CR)
- Binned in $N(b$-tags$)$, MET

**Other...**

**Multijet background**
Requires a data-driven estimate
1) Derive multi-jet template from multijet enriched region

2) Normalisation factor from fit using a multijet-sensitive variable

**Invariant mass template**

**Template without normalisation taken from QCD CR**

**Template scaled determined by fit**

**Multijet Normalisation**

<table>
<thead>
<tr>
<th></th>
<th>150 GeV &lt; $E_T^{miss}$ &lt; 200 GeV</th>
<th>200 GeV &lt; $E_T^{miss}$ &lt; 200 GeV</th>
</tr>
</thead>
<tbody>
<tr>
<td>1-btag</td>
<td>0.121</td>
<td>0.054</td>
</tr>
<tr>
<td>2-btag</td>
<td>0.130</td>
<td>2-btag</td>
</tr>
<tr>
<td>normalisation</td>
<td>0.068</td>
<td>0.036</td>
</tr>
<tr>
<td>relative uncertainty</td>
<td>0.199</td>
<td>0.081</td>
</tr>
</tbody>
</table>

**POSTFIT 150<MET<200 2 b-tag**
New requirement in resolved selection to suppress multijet background (poorly modelled by the dijet MC) at low values of the distribution.
Object-based MET significance separation power

- Met significance can help to identify and separate multijet background with respect to EW backgrounds and DM signals.

- Signal: DM signal: Z'-2HDM (mZ', mA) (400,300)
- Background: Dijet MC

More than 95% of dijet can be rejected while retaining a signal efficiency ~90%

**ATLAS**

*Work in Progress*

\( \sqrt{s} = 13 \text{ TeV} 
2 \text{ b-tag, Resolved}

**ATLAS**

Simulation Preliminary

\( \sqrt{s} = 13 \text{ TeV} 
2 \text{ b-tag, Resolved SR}

- Event-based \( E_T^{\text{miss}} \) Significance
- Object-based \( E_T^{\text{miss}} \) Significance

**ATLAS-CONF-2018-039**
VR TRACK JETS

- Canonical $b$-tagging: assumption of a single isolated $B$-hadron

- Mitigate loss of efficiency by redefining how single $b$-tags are obtained

- VR track jets: dynamic cone, $d_{ij} = \min\left(\frac{p_{T,j}^2}{\Delta R_{ij}^2}, \frac{p_{T,i}^2}{R_{\text{eff}}(\rho_{Ti})^2}\right)$

  arxiv 0903.0392
VR TRACK JETS OPTIMISATION

**Double Subjet B-Labelling Efficiency**

\[ \epsilon_{\text{Double subjet b-label}} = \frac{N(\text{Double subjet b-label} \mid \text{Higgs jet})}{N(\text{Higgs jet})} \]

jet in which the two b-hadrons are exclusively matched to the two leading \( p_T \) subjets

**SELECTION:** \( p_T > 250 \text{ GeV} \) and \( |\eta| < 2.0 \). The calorimeter jet mass is required to be between 76 GeV and 146 GeV which corresponds to a 90% signal selection efficiency to be consistent with \( m_{\text{Higgs}} \)