Searching for new physics at ATLAS
Rough guide to data analysis

Roland Jansky,
University of Geneva

HST17 - 11th July 2017
Overview

- LHC and ATLAS detector
- Physics objects
- Analysis – Boosted diboson search
- Summary Discussion
Particle physics – Why do we care?

• Ever wondered ..
  .. what am I and everything around me made of?
  .. how was our universe created?
  .. and what is it made of?

• We do!
  This is why we do particle physics.

Fraction of universe

- 95%
- 5%
- Roughly understood
- No clue!
Particle physics – Why do we care?

- What particles?

- These particles we know, and they are the building block of the 5%
Particle physics – Why do we care?

• What else could there be?

• Are these the 95%?
Physics analysis needs:
- Leptons
  - Muon, electron, tau
- Photons
- Hadronic jets
- Missing transverse energy
Energy Frontier: Jets

- Newly opened energy regime: $\sqrt{s} = 13$ TeV.

m=5.2 TeV dijet event
$jet_{1/2}p_T=2.5/2.4$ TeV

something new?

proton  X  proton

11/07/2017  R. Jansky - Searching for new physics at ATLAS: Rough guide to data analysis
• Nowadays finding particle tracks is like this ..
• Nowadays finding particle tracks is like this ..

• But without the numbers!
1. Register measurements (called clusters).
1. Register measurements (called clusters).
2. Associate clusters to particles’ tracks.
1. Register measurements (called clusters).
2. Associate clusters to particles’ tracks.
3. Fit particles trajectory.
1. Create measurements (called clusters).
2. Associate clusters to particles’ tracks.
3. Fit particles trajectory.

How it looks in reality
Track reconstruction – Step by step

1. Create measurements (called clusters).
2. Associate clusters to particles’ tracks.
3. Fit particles trajectory.
Prerequisites for analysis

1. Collect data: Detector, trigger, DAQ

2. Reconstruction of physics objects

3. Simulation: Generate events, detector simulation
Boosted diboson search strategy

- Heavy (>1 TeV) resonances to pairs of vector bosons (V=W/Z) predicted by several extensions of the SM.
- $V \rightarrow$ quark-pair decays most abundant. 
  ➔ Great probe for new physics!
- Mass of jet can identify initiating particle.

$m=?$
Boosted diboson search strategy

R-S graviton

$X \rightarrow V V$

$G^*$

$Z/W$

$m=?$

11/07/2017
Boosted diboson search strategy

- Look for bump in steeply falling invariant mass distribution.

Diagram:
- Y-axis: Events
- X-axis: Invariant mass
- Graph shows a background line and a bump indicating a significant deviation from background.
Main background:

- Standard model processes that can give the same 2-jet signature
“Distinguishing the signal from the noise requires both scientific knowledge and self-knowledge: the serenity to accept the things we cannot predict, the courage to predict the things we can, and the wisdom to know the difference.”

— Nate Silver, *The Signal and the Noise: Why So Many Predictions Fail - But Some Don't*
How to measure a cross-section

Cross section \[ \sigma = \frac{N}{\mathcal{L}} \rightarrow \sigma = \frac{N_{obs} - N_{bkg}}{\mathcal{L} \cdot \epsilon \cdot A \cdot B} \]

N(obs) = Observed number of events

N(bkg) = Estimated number of background

\( \mathcal{L} = \) Integrated luminosity

\( \epsilon = \) efficiency

A = acceptance

B = Branching ratio
How to measure a cross-section

- Cross section
  \[ \sigma = \frac{N_{\text{obs}} - N_{\text{bkg}}}{\mathcal{L} \cdot \epsilon \cdot A \cdot B} \]

  - \( N(\text{obs}) \) Direct from data
  - \( N(\text{bkg}) \) (from data and MC, most critical part of analysis)
  - \( \mathcal{L} \) (Someone else calculates this!)
  - \( \epsilon \) = efficiency (from Monte Carlo)
  - \( A \) = acceptance (from Monte Carlo)
  - \( B \) = Branching ratio (from Particle data group)
Arise from stochastic fluctuations arising from the fact that a measurement is based on a finite set of observations.

Repeated measurements will give a set of observations that will differ from each other.

Statistical uncertainty is a measure of this variation.

Poisson fluctuations associated with random variations in the system one is examining.
Systematic Uncertainties

- Arise from uncertainties associated with the measurement apparatus
- What are the assumptions underlying the measurement?
  - How accurate is the Monte Carlo Simulation?
  - Models for the signal and the background
  - E.g. acceptance, model parameters
  - What can we think of that has the potential to affect our measurement?
Cross section $\sigma = \frac{N_{\text{obs}} - N_{\text{bkg}}}{\mathcal{L} \cdot \epsilon \cdot A \cdot \mathcal{B}}$

- $N(\text{obs})$ Statistical uncertainty
- $N(\text{bkg})$ Systematic uncertainty
- $\mathcal{L}$ Systematic uncertainty
- $\epsilon$ Systematic uncertainty
- $A$ Systematic uncertainty
- $B$ Systematic uncertainty
Optimisation

\[ \sigma = \frac{N_{\text{obs}} - N_{\text{bkg}}}{L \cdot \epsilon \cdot A \cdot B} \]

- Minimise the uncertainty on \( \sigma \).
- Maximise probability for signal detection, minimise probability for arriving at a fake signal detection.
- High signal to background : \( N(\text{obs}) \gg N(\text{bkd}) \)
- High signal efficiency \( \epsilon A \)
- Reliable, robust method to determine \( N(\text{bkg}) \).
- Most important is the measurement of the uncertainty on \( N(\text{bkg}) \)
- Use Monte Carlo to help decide selection criteria that attempt to minimise the uncertainty on \( \sigma \) or significance of a discovery.
• Try to maximize fraction of signal events versus background events.
• One example is the use of kinematic cuts.

Cut on separation of two jets in rapidity
• Try to maximize fraction of signal events versus background events.
• One example is the use of kinematic cuts.

Cut on momentum asymmetry between two jets
• Utilize different properties of jets from **W/Z-bosons** with respect to the background to “tag” them.

**Cut on jet substructure** (calculated from energy distributions inside fatjet)
• Utilize different properties of jets from W/Z-bosons with respect to the background to "tag" them.

Cut on number of tracks in jet

• Selection on jet mass: require to be at W/Z mass!
• After fixing selections of analysis, calculate expected signal efficiency and yield.
Control regions

- Before looking at the data, validate analysis in control regions (e.g. mass sidebands).
- Check that background is smoothly falling and not sculpted by selections.
1. Finally, look at signal region

2. Fit background

3. Check for difference between fit and data.

**Result from run 1!**
Statistical treatment

- Estimate of p-value/significance of observed events, assuming probability density for random variable
- Assume: $N_0$ follows Poisson distribution

Poisson probability: $\alpha = \sum_{n=N_0}^{\infty} \frac{\exp \left( -N_b \right) (N_b)^n}{n!}$. 
Statistical treatment

- Estimate of p-value/significance of observed signal assumption
- Assumption of Poisson density
- Poisson probability: $\alpha = \sum_{n=N_0} \frac{\exp(-N_b) (N_b)^n}{n!}$

**Does not consider uncertainties or possible other intervals where to measure $N_0$**

$\Rightarrow$ local significance versus global significance
1-2\sigma
Statistical treatment

3σ
Statistical treatment

$5\sigma$
More Results!

**ATLAS** Preliminary
\( \sqrt{s} = 13 \text{ TeV, } 15.5 \text{ fb}^{-1} \)

- **Events / 0.1 TeV**
- **WZ selection**

- **Data 2015+2016**
- **Fit bkg estimation**
- **HVT Model A \( m=1.5 \text{ TeV} \)**
- **HVT Model A \( m=2.4 \text{ TeV} \)**
- **Fit exp. stats error**

---

**Pull**

**\( M_{JJ} \text{ [TeV]} \)**

20/07/2016  Roland Jansky - Searching for new physics at ATLAS: Rough guide to data analysis
ATLAS Preliminary
$\sqrt{s} = 13$ TeV, 36.7 fb\(^{-1}\)
WZ SR
$\chi^2$/DOF = 8.1/9

Data
Fit
Fit + HVT model B $m=1.5$ TeV
Fit + HVT model B $m=2.4$ TeV

Significance

$m_{JJ}$ [TeV]
Essentially thin jets \( \tau^+ \rightarrow \pi^+ \nu_\tau \)

\( \tau^+ \rightarrow \pi^+ \pi^+ \pi^- \nu_\tau \)