Jet Clustering Techniques for New Higgs Boson Searches in Hadronic Final States

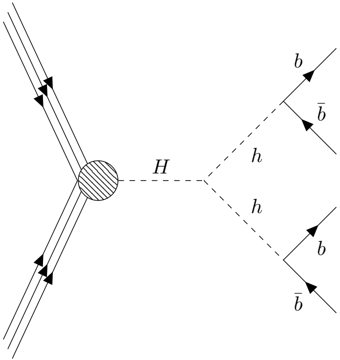
Shubhani Jain

University of Southampton

ICHEP 2022, Bologna, Italy

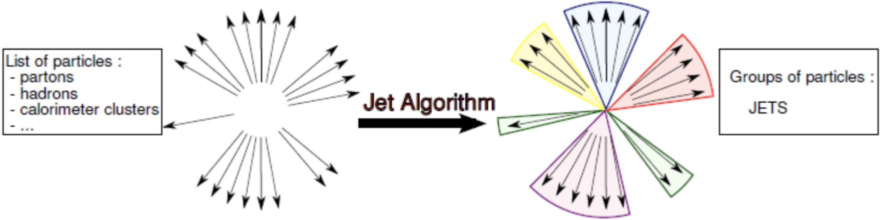
Motivation

- Ultimate motivation of our study is to look for signs of physics beyond the Standard Model (BSM).
- We investigate whether different jet clustering techniques might be more or less suited to particular final states of interest.



Jet Clustering Algorithms

- In order to extract proper physics from hadronic sprays found in particle detectors, algorithms are used to characterise the detected radiation into distinguishable objects, the aforementioned jets.
- There are two main classes of jet algorithms in use: sequential recombination algorithms and cone algorithms.



Credit: http://www.kip.uni-heidelberg.de/atlas/seminars/SS2009 JC/jet algorithms.pdf

Sequential Jet Clustering Algorithms

- We only focus on sequential recombination algorithms in our study.
- These algorithms take a similar form, descending from the generalised k_T algorithm:

$$d_{ij} = min(p_{Ti}^{2a}, p_{Tj}^{2a})\frac{\Delta R_{ij}^2}{R^2} \qquad d_{iB} = p_{Ti}^{2a}$$

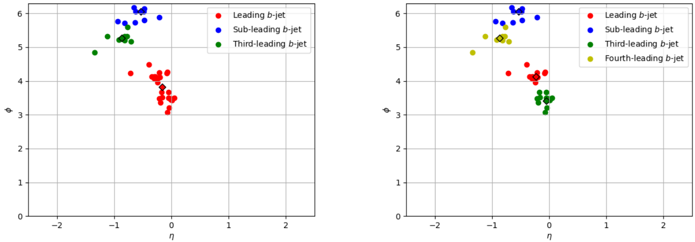
- The algorithm works by finding the minimum d_{min} of all the d_{ij}, d_{iB} and then if d_{min} is a d_{ij}, combine i and j and repeat the process or if d_{min} is a d_{iB}, then i is declared a jet and removed from the list.
- The algorithms currently used at the LHC are the anti-k_T one and the Cambridge/Aachen (CA) one.

Variable-R Algorithm

- The more recent variation of the standard jet clustering algorithms is the so-called Variable-R.

$$d_{ij} = min(p_{Ti}^{2a}, p_{Tj}^{2a})\Delta R_{ij}^2 \qquad d_{iB} = p_{Ti}^{2a} R_{eff}(p_{Ti})^2$$

where $R_{eff}(p_{Ti}) = \frac{\rho}{p_T}$ and ρ is a dimensionful input parameter. There are other two parameters such as $R_{max/min}$, which are cut offs for the maximum and minimum allowed R_{eff}. These parameters can be optimised for better results.



Selection of Suitable Benchmark

- We first select a suitable set of parameters in the 2HDM Type-II framework for our model.
- We work with two sample Benchmark Points (BPs) that we call BP1 and BP2:

	m_H (GeV)	m_h (GeV)
BP1	700	125
BP2	125	60

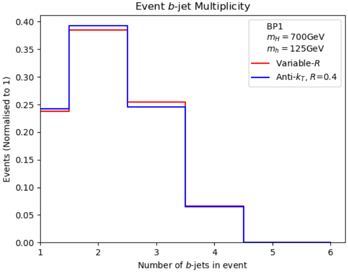
  where m_H is heavy Higgs and m_h is light Higgs boson. For both the BPs the 125 GeV is SM like Higgs boson.
- We have tested these benchmarks against theoretical and experimental constraints by using 2HDMCalculator interfaced with HiggsBounds and HiggsSignals and flavour constraints using SuperISO.

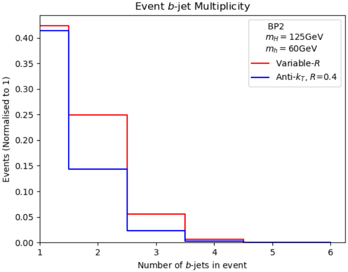
Simulation Details and Cutflow

- Generate samples of signal events of $\mathcal{O}(10^5)$ with $\sqrt{s} = 13$ TeV for the process $gg \to H \to hh \to b\bar{b}b\bar{b}$ using MadGraph5.
- Shower and hadronise parton level events using Pythia8.
- Apply detector simulation using Delphes to output eflow objects.
- Perform jet reconstruction on the eflow objects in FastJet using MadAnalysis5.
- Remove jets with $p_T < 50$ GeV (BP1) / 20 GeV (BP2)
- Implement a simplified MC informed b-tagger on clustered jets after cuts have been applied.

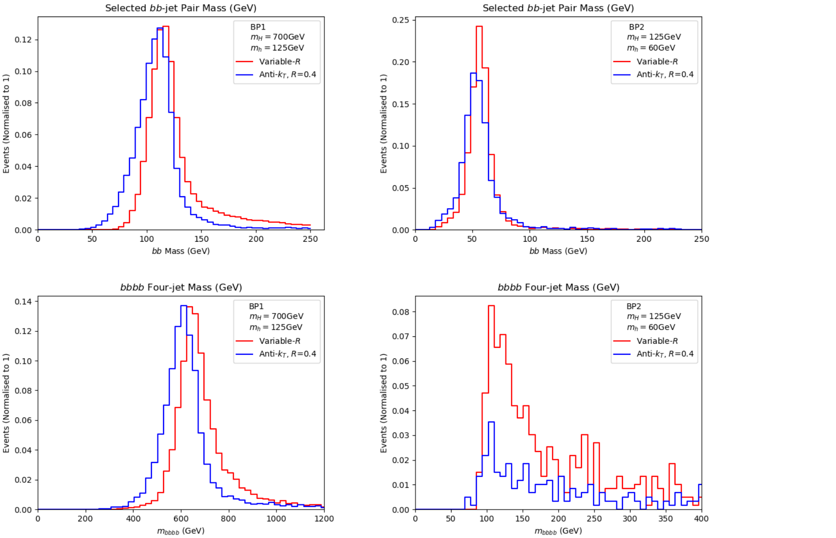
Results

- We start by investigating the b-jet multiplicity plots for variable versus a fixed-R analysis. We choose a value of $R = 0.4$ for fixed-R and $\rho = 100$ for Variable-R. We have used Anti-k_t clustering algorithm for both the cases.



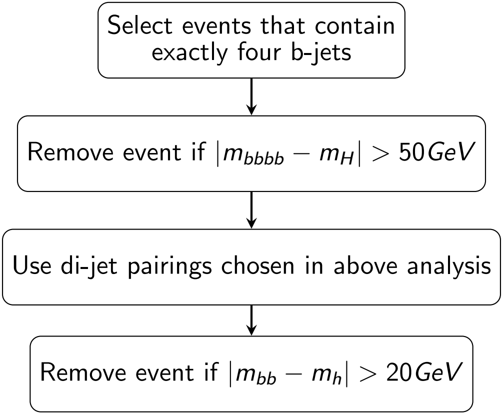


- We look at the invariant mass of pairs of clustered *b*-jets (dijets) and fourjets, in order to reconstruct the masses of the resonance from which they originated.



Signal-to-Background Analysis

- As a final exercise, we perform a calculation of the signal-to-background rates to compare the various jet reconstruction procedures used in our study. To calculate the rates, we generate and analyse $pp \to b\bar{b}b\bar{b}$, $pp \to t\bar{t}$ and $pp \to Zb\bar{b}$ background processes.



- We calculate the significance rates (Σ) for two values of (integrated) luminosity:

$$\Sigma = \frac{N(S)}{\sqrt{N(B_{b\bar{b}b\bar{b}}) + N(B_{t\bar{t}}) + N(B_{Zb\bar{b}})}}$$

For $\mathcal{L} = 140\ fb^{-1}$, the final Σ values are:

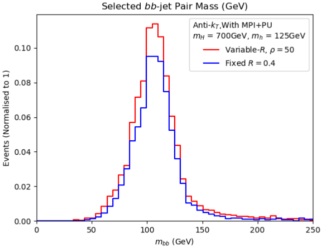
	variable-R	$R = 0.4$
BP1	1.881	1.366
BP2	3.707	1.984

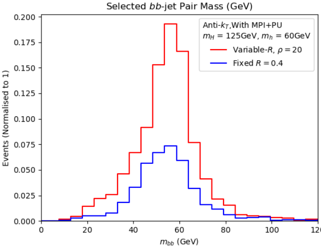
For $\mathcal{L} = 300\ fb^{-1}$, the final Σ values are:

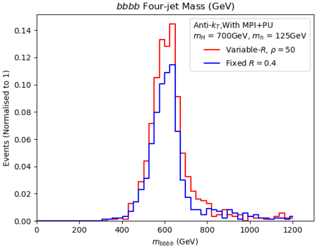
	variable-R	$R = 0.4$
BP1	2.753	2.000
BP2	5.426	2.905

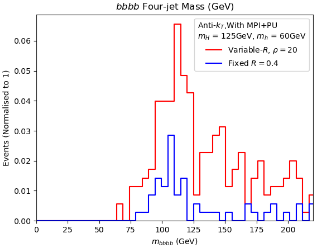
Pile-Up

- For completeness, we also present the PileUp (PU) plots for both reconstruction procedure.









Conclusion and Outlook

- We have demonstrated that the potential scope of the LHC experiments (from a theoretical perspective) in accessing BSM Higgs signals are suppressed with current jet reconstruction parameters.
- Variable-R jet clustering can outperform fixed-R implementations currently in use.
- The results can be applied to any BSM models with four-b final states, e.g., $H \to AA \to b\bar{b}b\bar{b}$.
- We are currently working on using fixed-R algorithm and Variable-R alogrithm to analyse fatjets and then visualise these fatjets using ML techniques such as image recognition to distinguish between signal and backgrounds jets.
- More details of our study can be found here:arXiv:2008.02499, epjc/s10052-022-10314-z.

Thank you for listening!