

# Jet SIFT-ing for New Physics in a Hidden Valley

James Floyd, Jonathan Mellenthin, Camryn Sanders



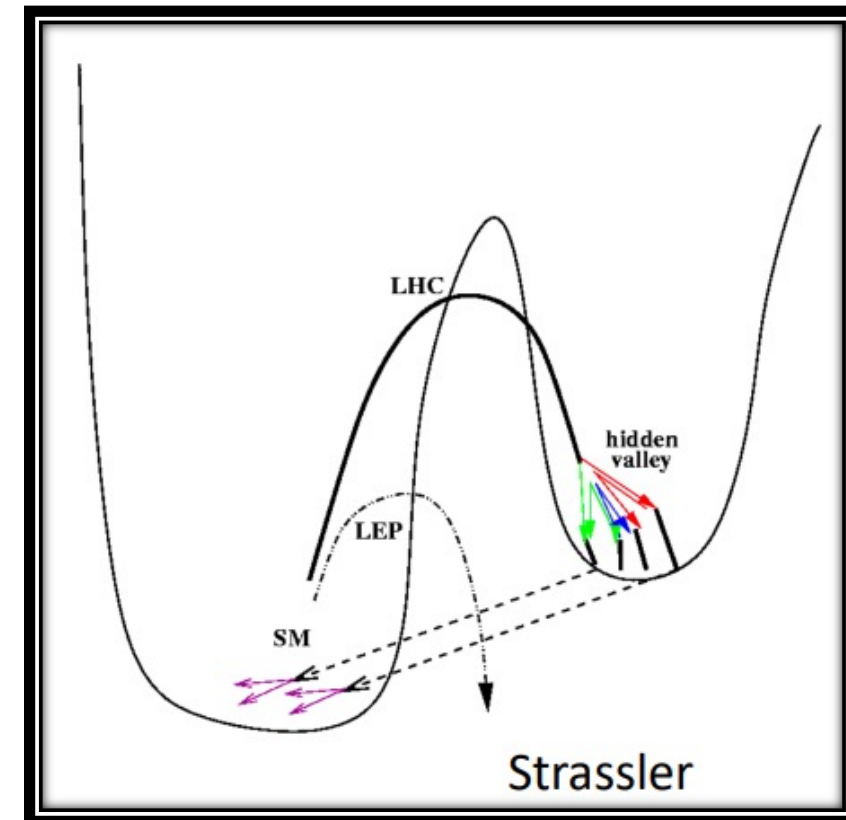
In collaboration with Shepherd and Walker

# Hidden Valley Background

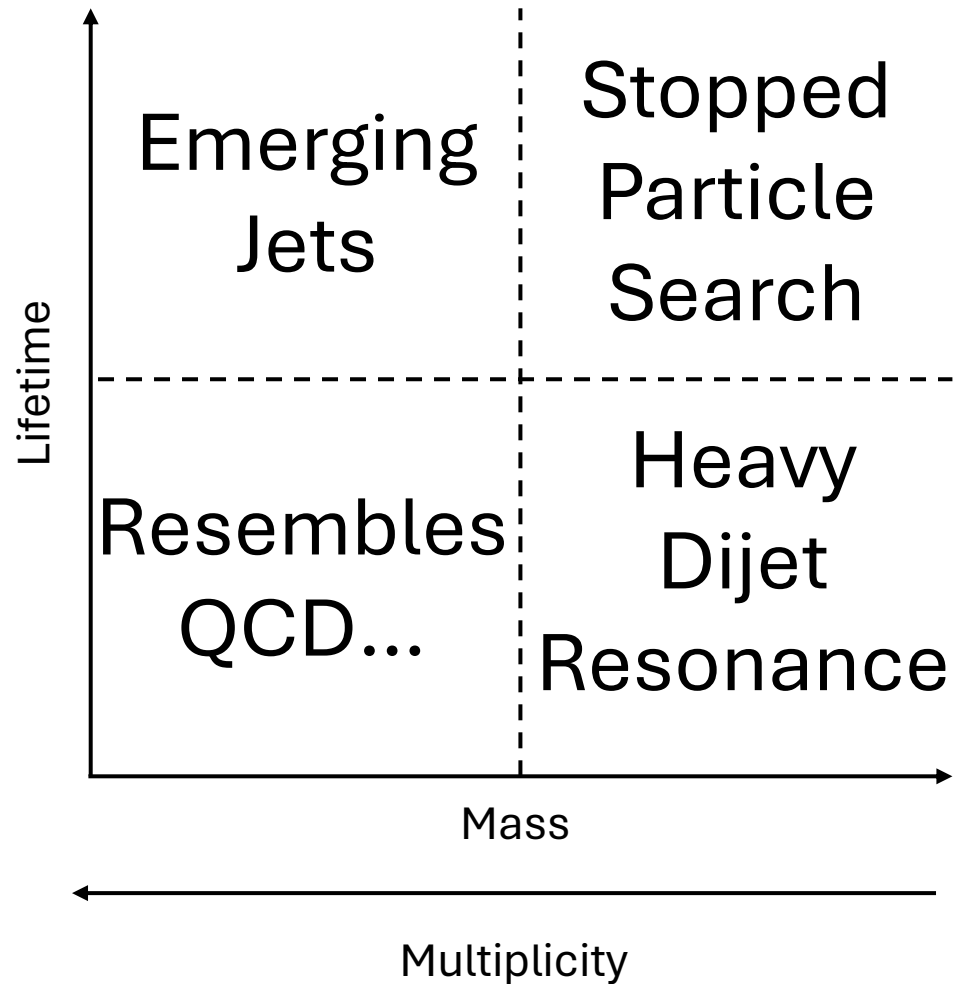
The Hidden Valley Scenarios were described by Strassler and Zurek leading up to the start of collisions at the LHC (hep-ph/0604261)

- ❑ Characterized by new light physics that is weakly coupled to the Standard Model
- ❑ A heavy intermediary ( $Z'$ ) presents a high energy barrier to access the new sector
- ❑ Strong dynamics & confinement are expected
- ❑ A mass gap allows decays back to the Standard Model

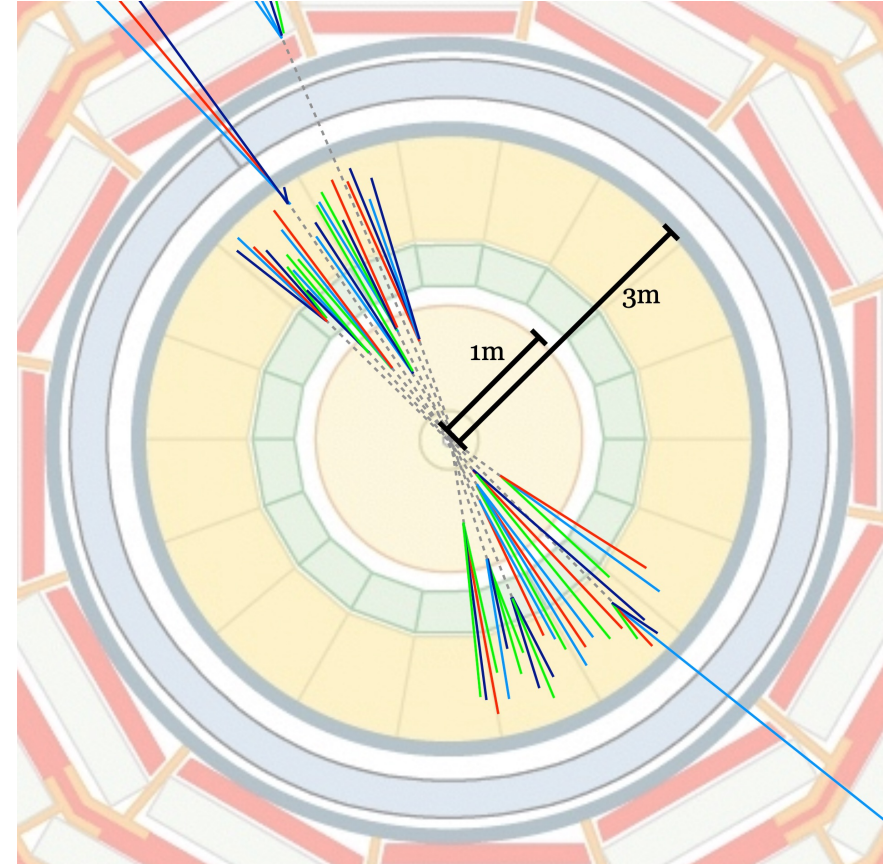
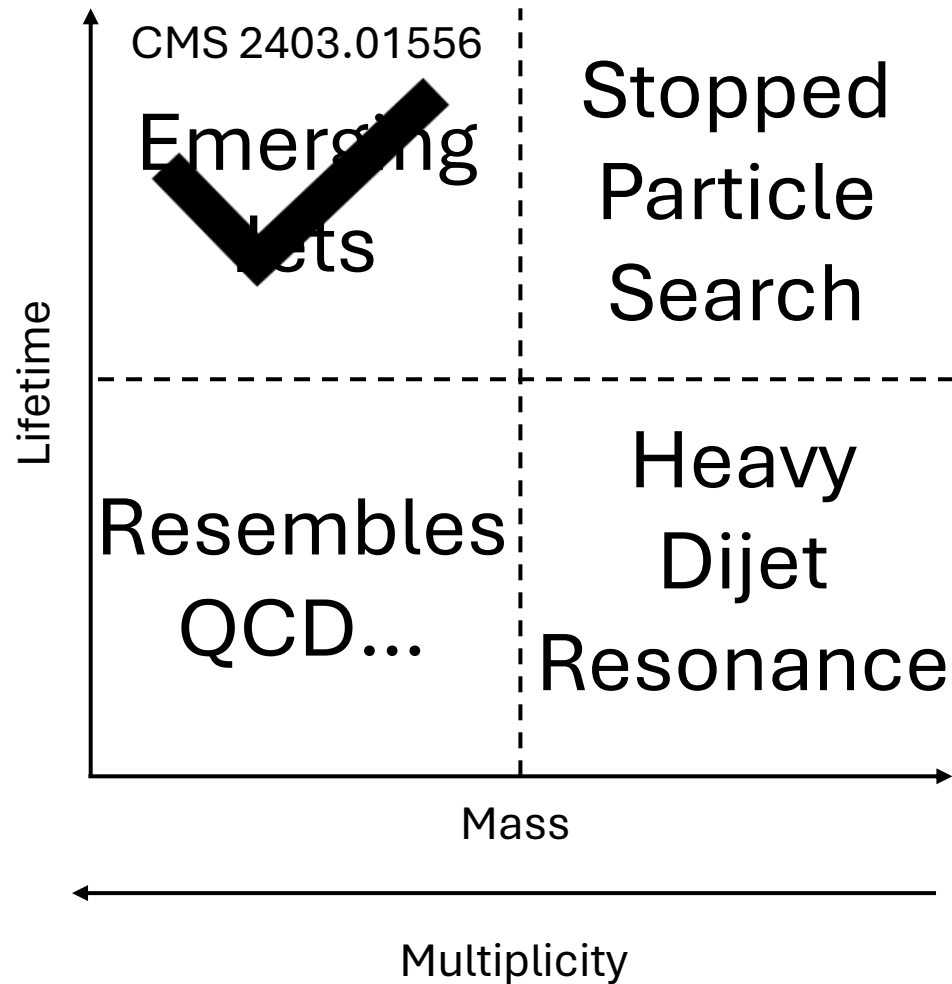
“ A unexpected place ...  
... of beauty and abundance ...  
... discovered only after a long climb ...”



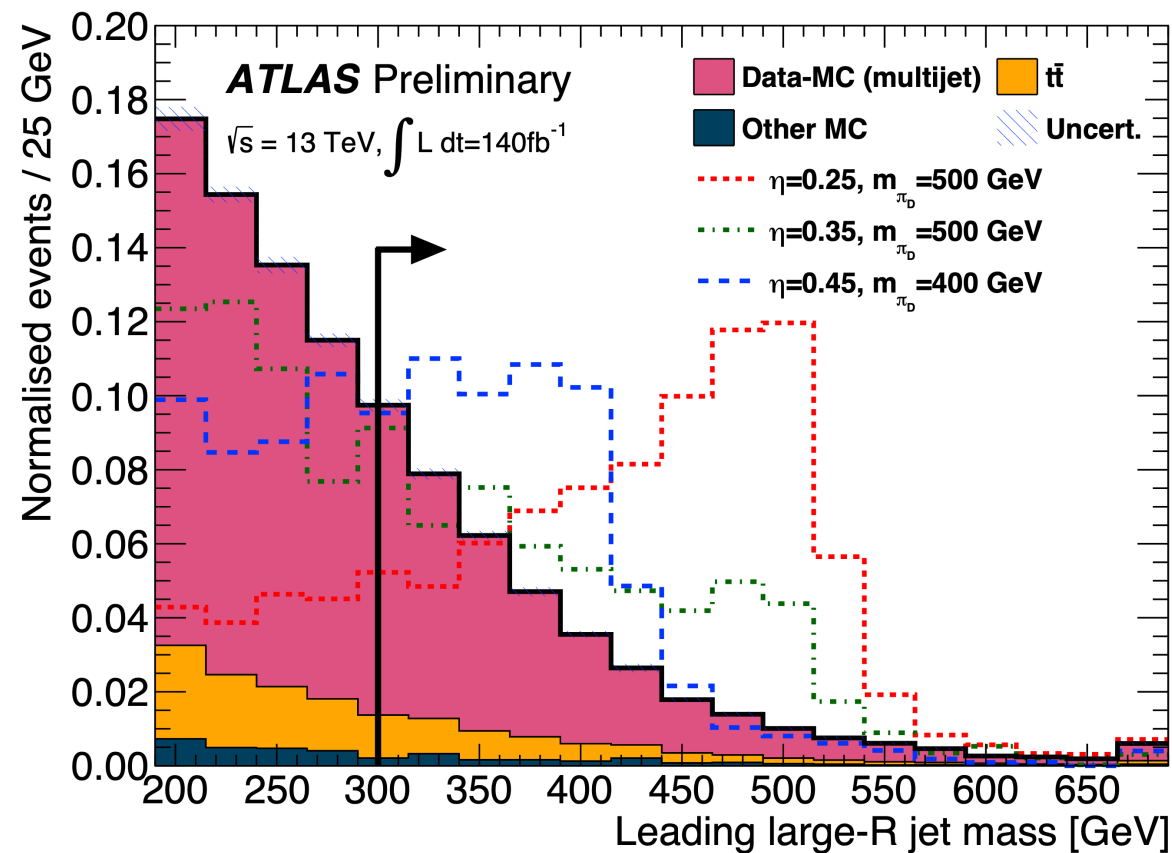
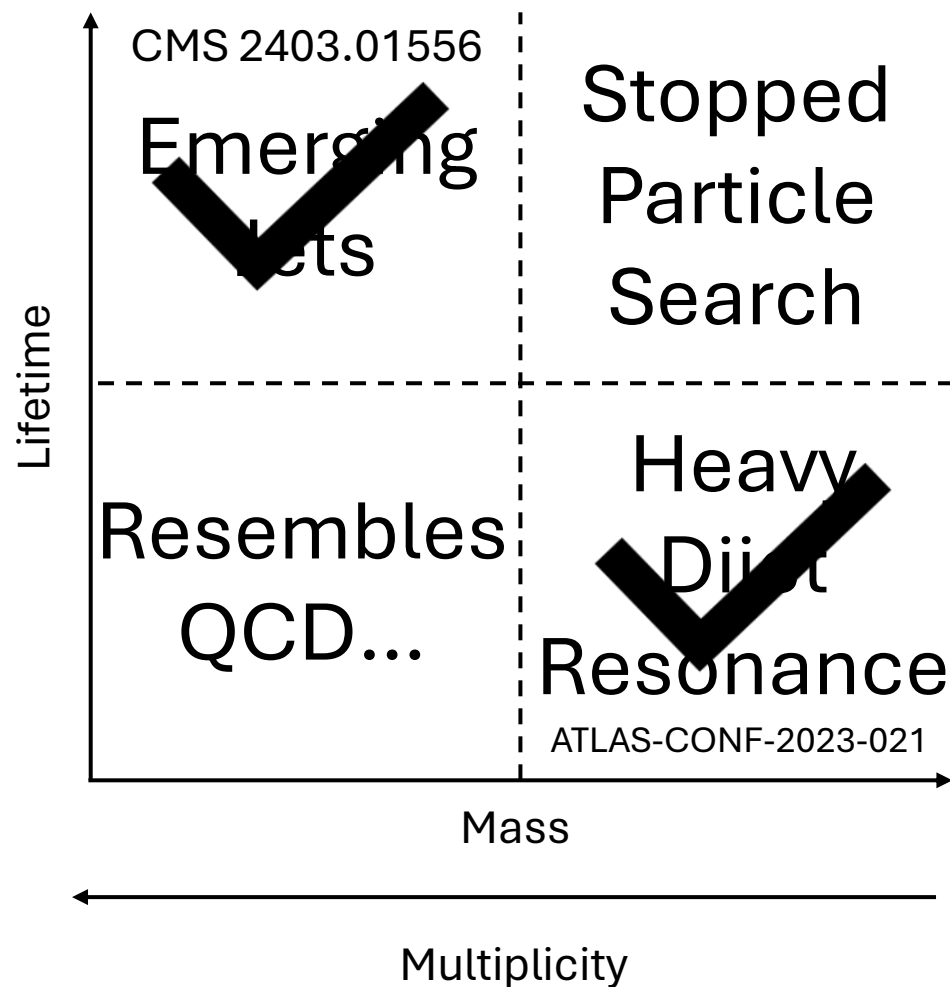
# Hidden Valley Signatures



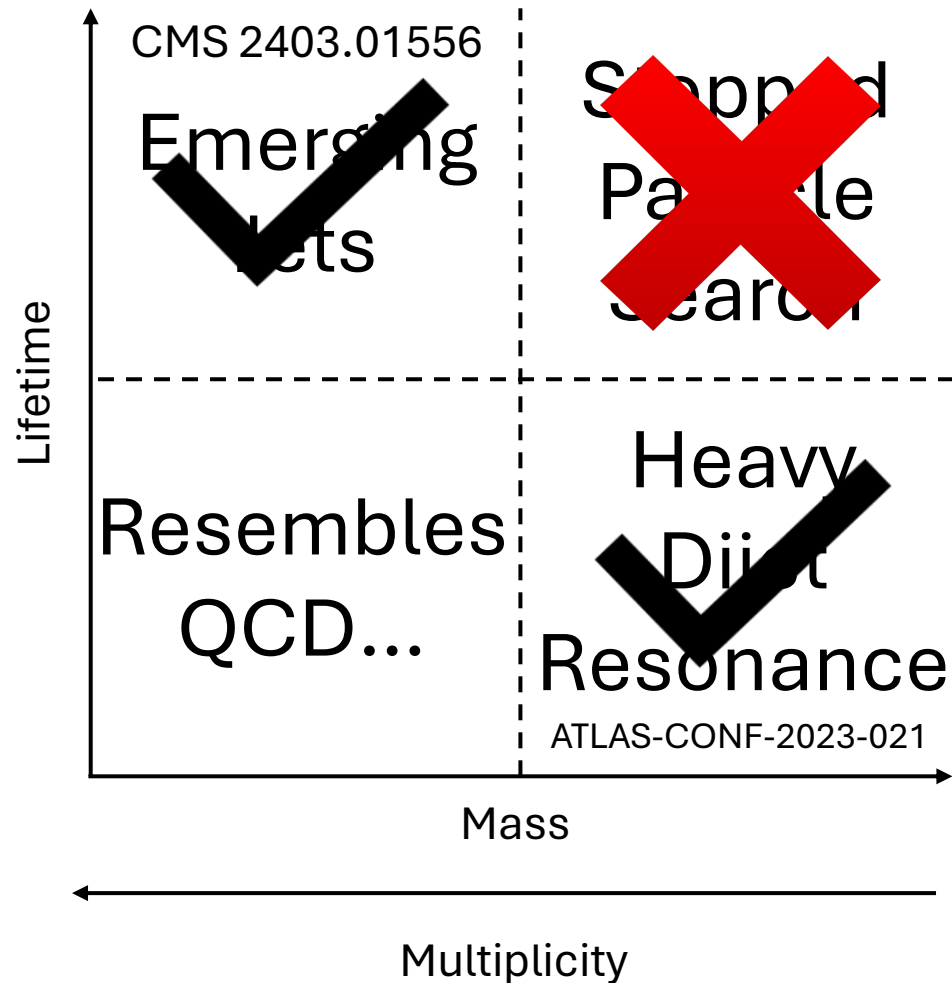
# Hidden Valley Signatures



# Hidden Valley Signatures

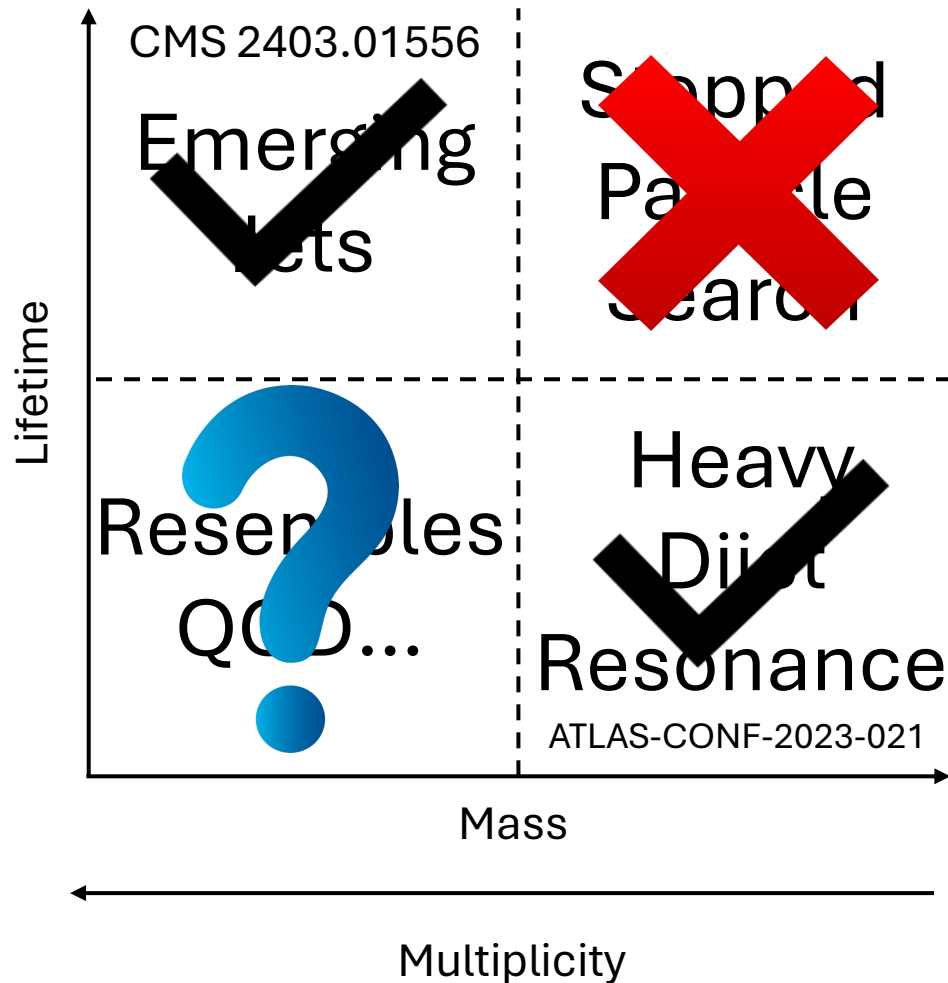


# Hidden Valley Signatures



- Large mass with long lifetime means very small couplings
- Negligible particle production at LHC

# Hidden Valley Signatures



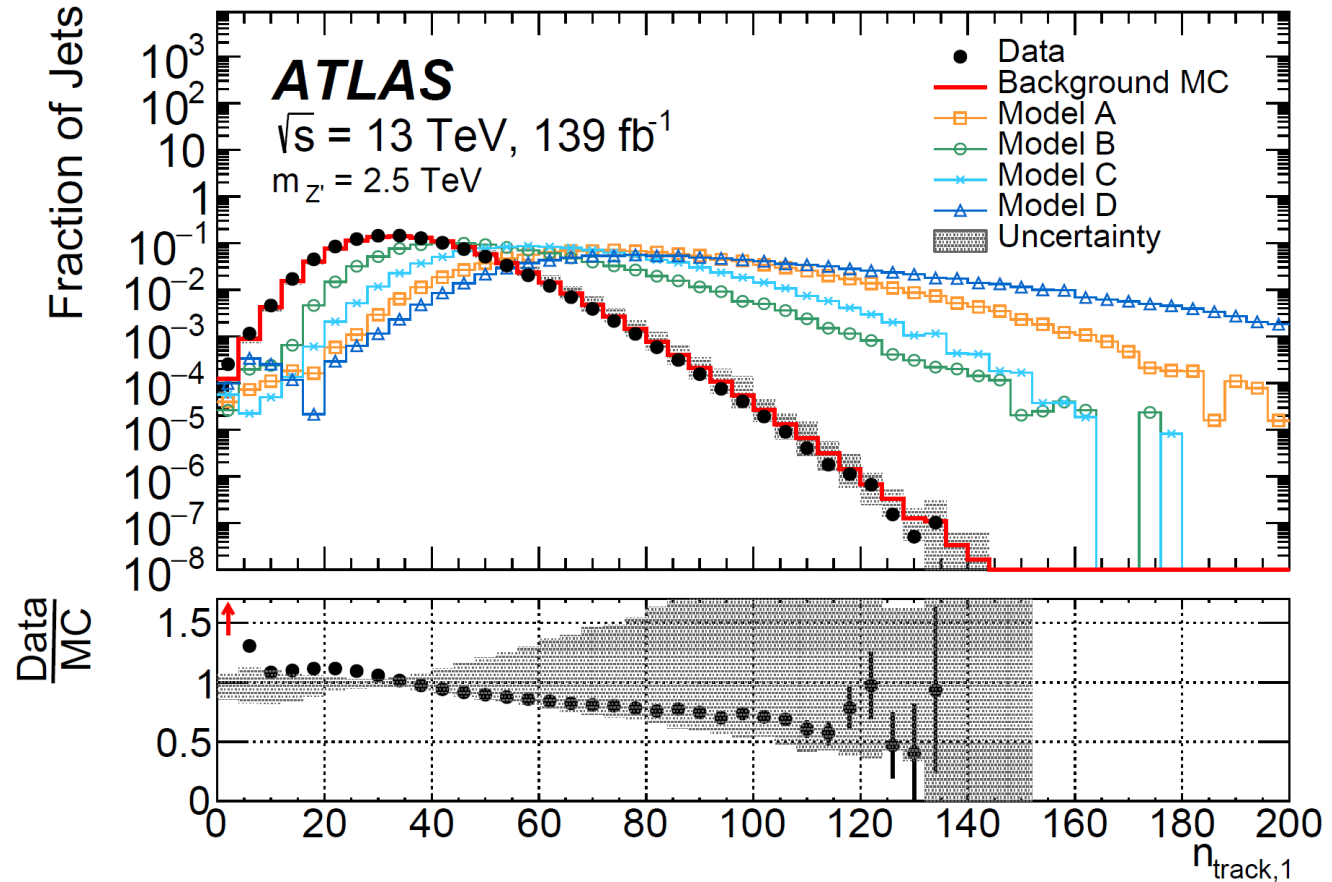
## What we want to focus on

- Resembles QCD -
  - This explains the absence of any searches in this area so far.
  - Our goal is to discern just how closely it resembles QCD and distinguish it
- Combinatoric background –
  - Due to the large number of possible pairings, reconstructing the physical dark pion mass is very challenging

# ATLAS Search for Dark Jets

arXiv:2311.03944

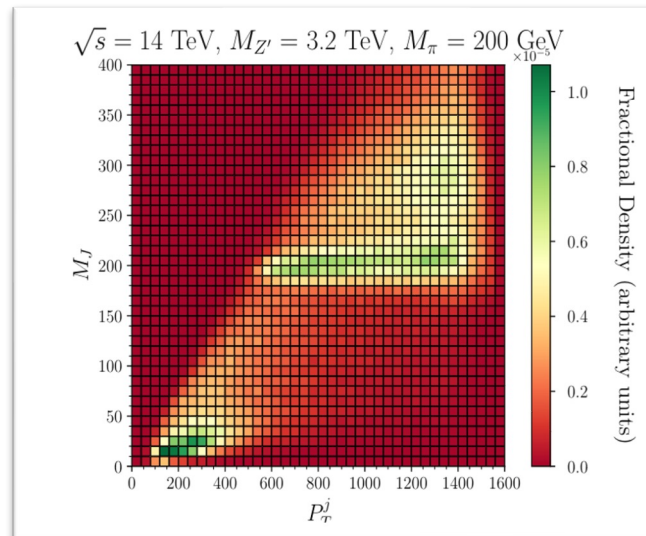
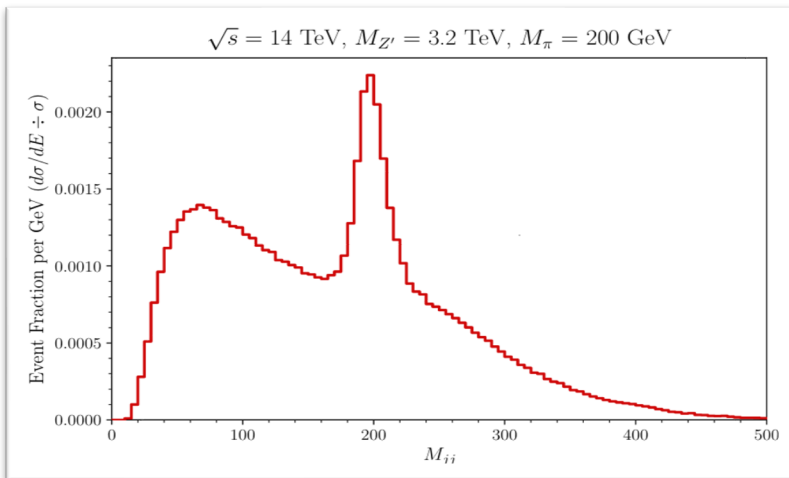
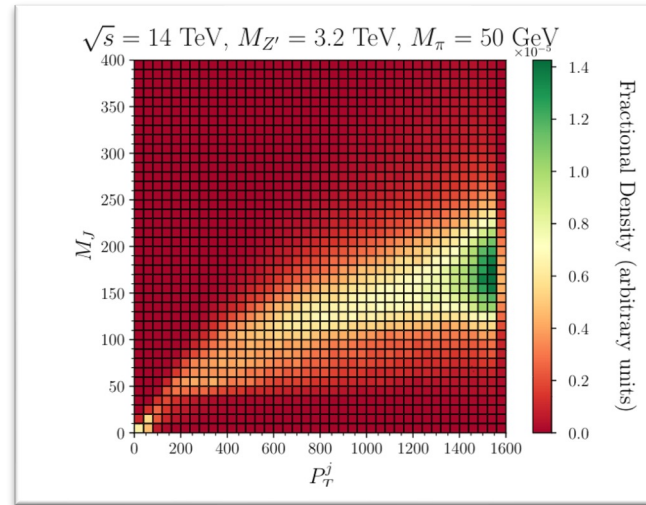
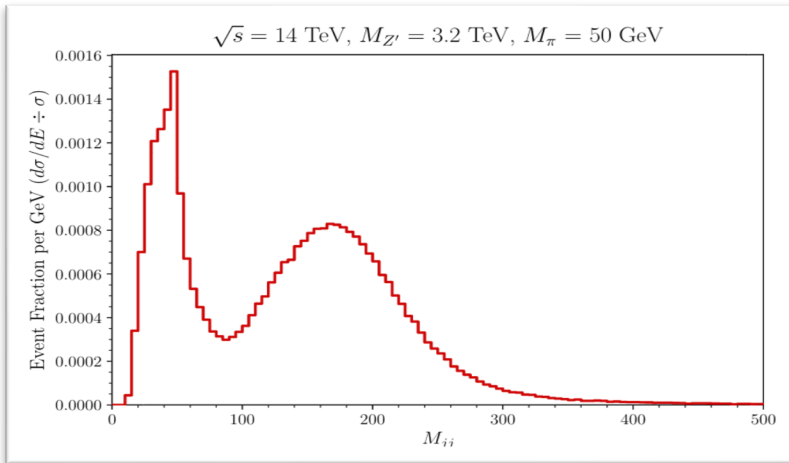
- ❑ Search for Z-prime production with decays into 'dark quarks' at ATLAS
- ❑ Selects signal events by anomalously large number of tracks in jets at a given  $P_T$
- ❑ Aims to reconstruct Z' mass, but not to see the dark pions directly





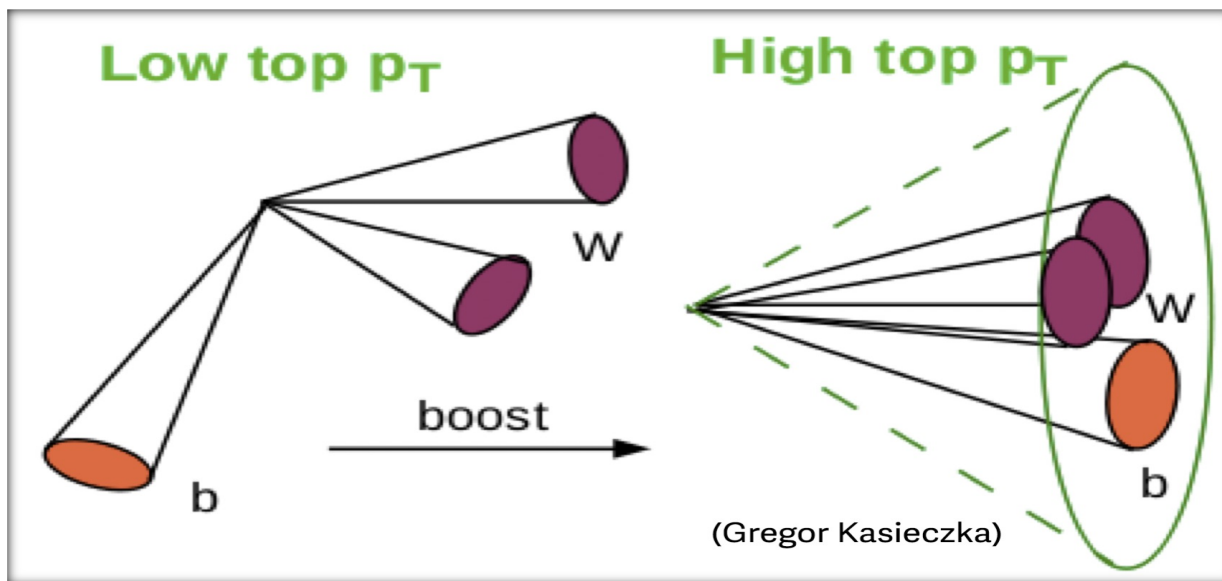
# Dijet Resonance Techniques

*A la* Strassler arXiv:0806.2385



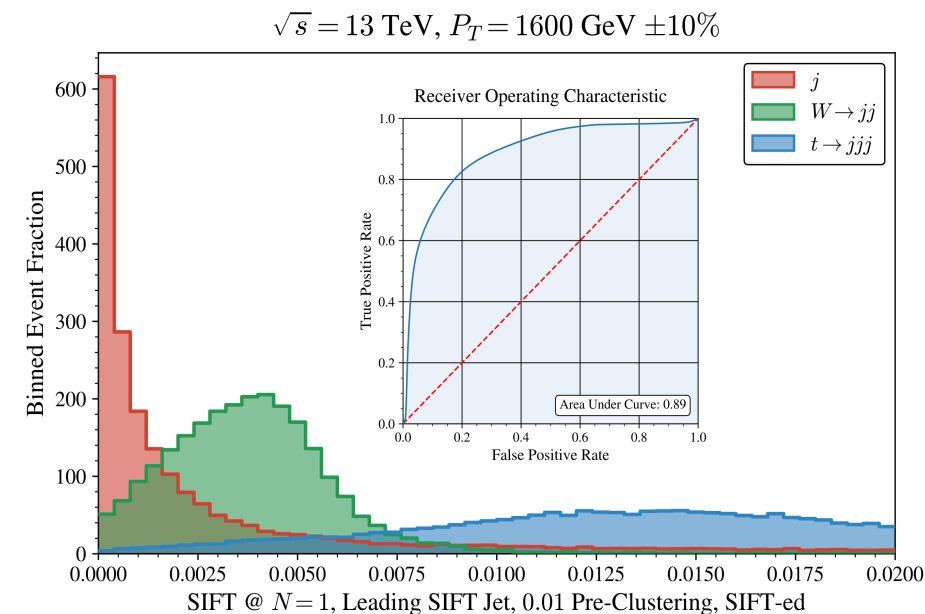
- Assumed that dark pions decay into bottom quarks
  - Effectively background free
- Reconstructed resonances in dijet masses or monojet masses
- This reconstruction works much better for higher masses

# SIFT: Scale-Invariant Filtered Tree



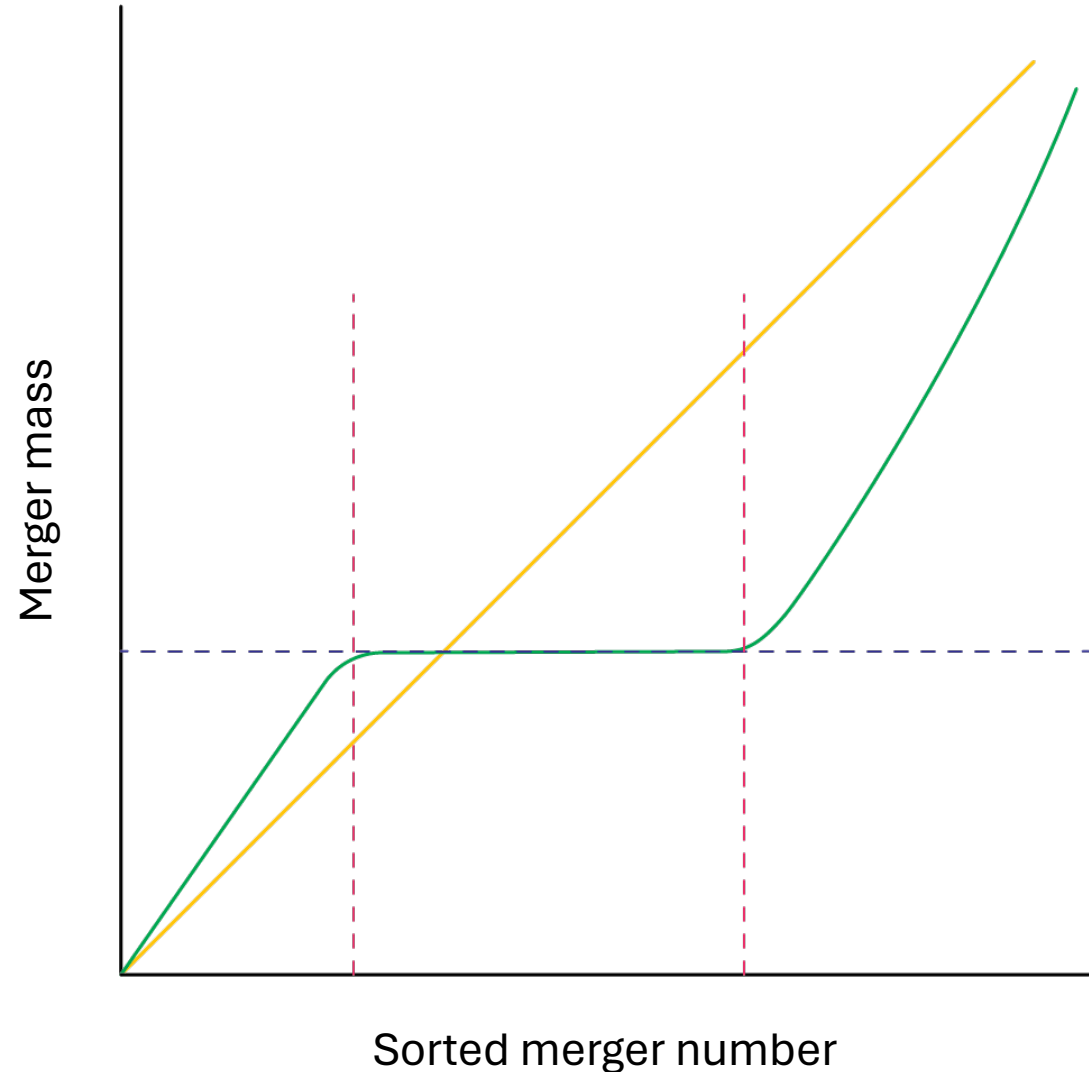
$$\delta_{AB} \equiv \frac{\Delta M_{AB}^2}{E_{TA}^2 + E_{TB}^2}$$

- ❑ A radius-free jet clustering algorithm
- ❑ Mutually hard prongs cluster last
- ❑ Tags substructure while clustering
- ❑ Well-defined combinatoric slice
- ❑ One method for low & high boost



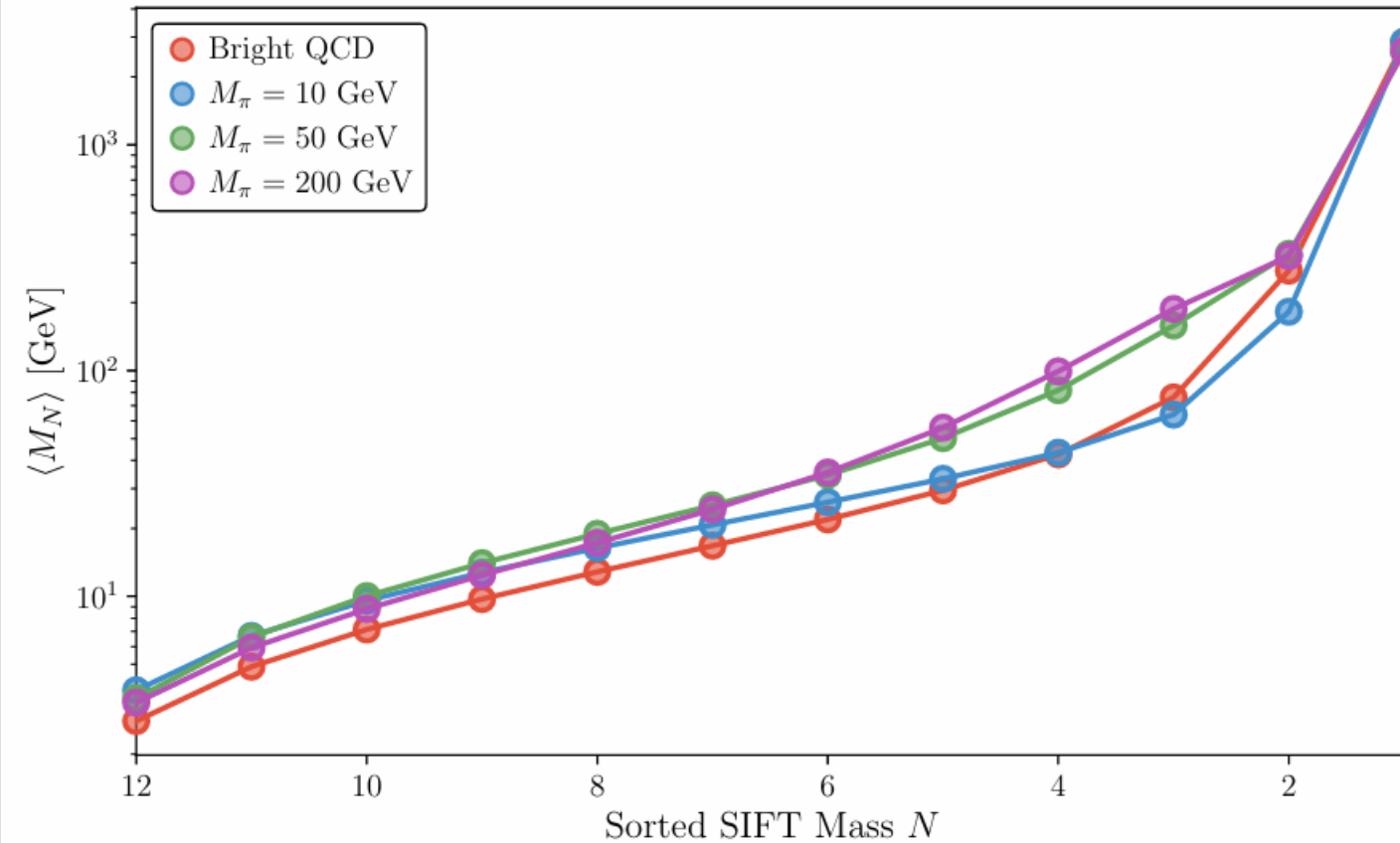
# Plateau Investigation Of New Scales: PIONS

- ❑ Cartoon of a log plot of jet algorithm merger masses, sorted by mass
- ❑ **Yellow** line shows a typical QCD showering
- ❑ **Green** shows a Hidden Valley plus QCD showering
- ❑ **Red** upright lines are separated by the hypothetical plateau length  $N$ 
  - This length is a parameter defining the variable
- ❑ Slope of the **Blue** line is the variable we want to search in
  - We calculate the rms of discretized derivatives between mergers to yield the variable  $\Pi_N$



# Exploring Merger Masses in Simulation

$\sqrt{s} = 14 \text{ TeV}$

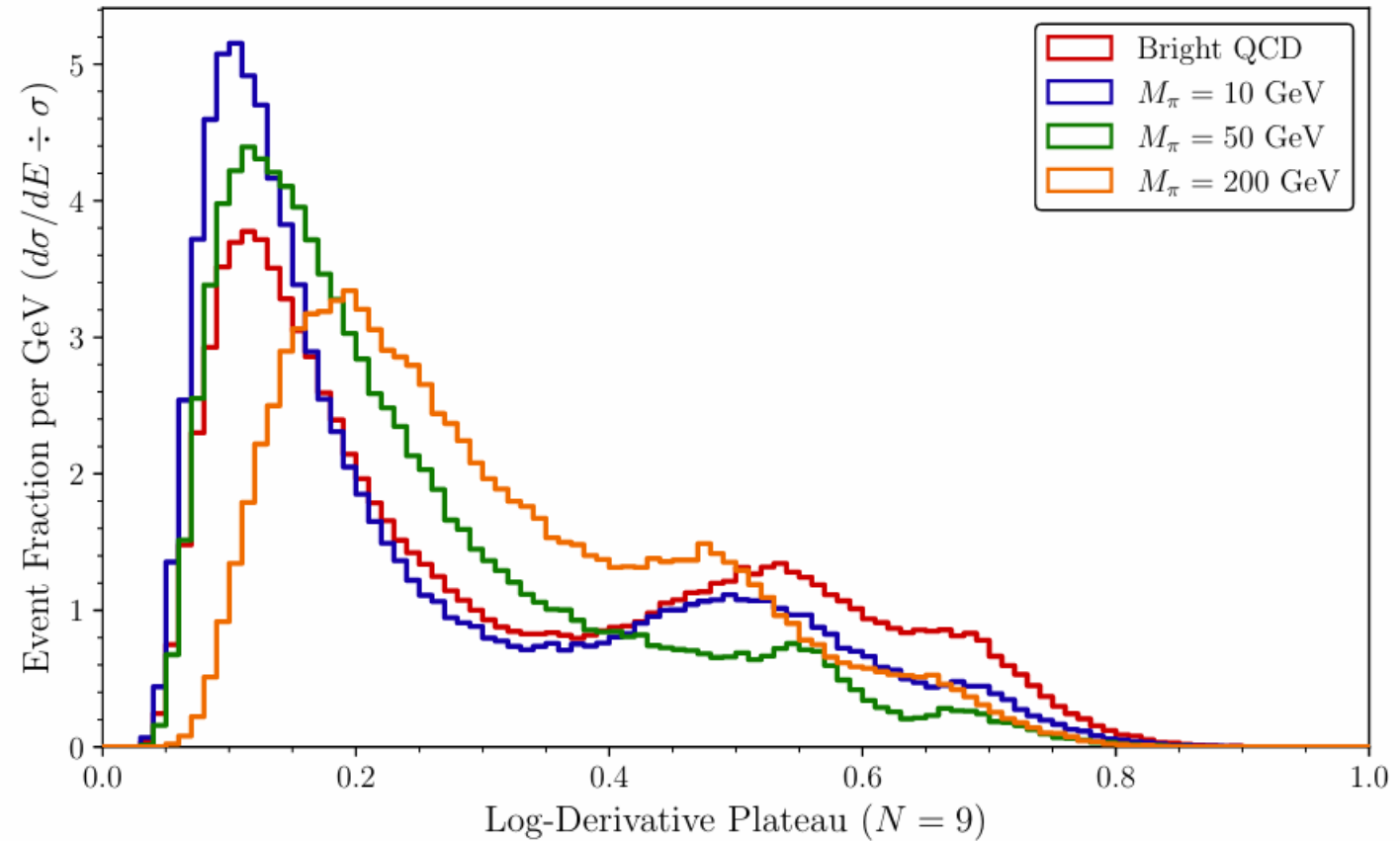


- This is the sorted merger mass plot averaged over a large number of events.
- Note flattening in blue curve relative to red

# PIONS distribution

- ❑ Event-by-event  $\Pi_9$  for different showering hypotheses
- ❑ A flat plateau of length  $N$  in the merger mass plot gives low values of  $\Pi_N$

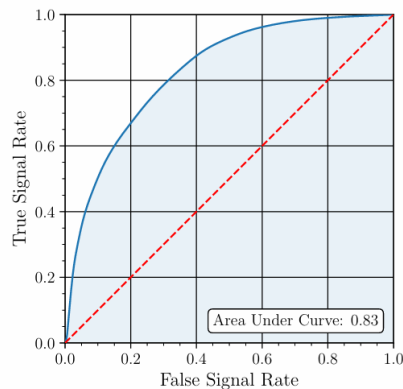
$\sqrt{s} = 14 \text{ TeV}, M_{Z'} = 3.2 \text{ TeV}$



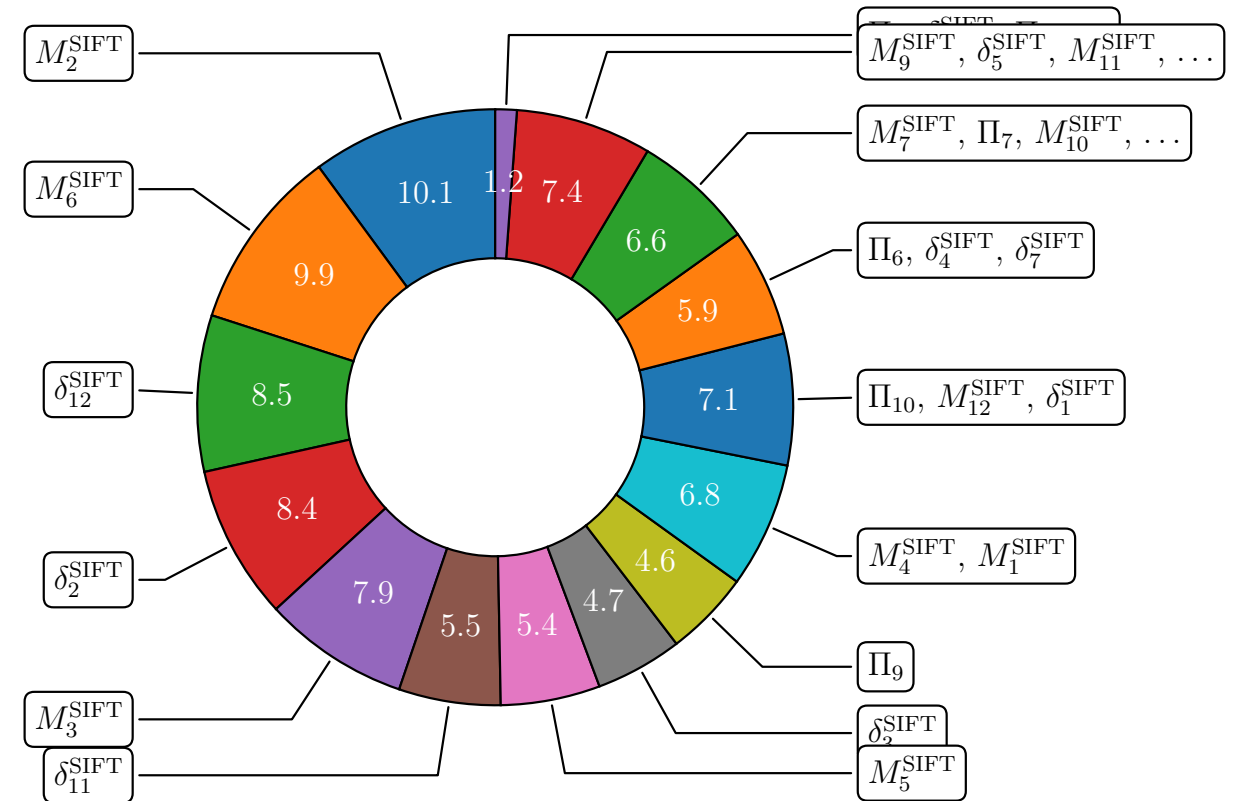
# BDT Results $M_{\pi} = 10$ GeV – Our Variables

- ❑ Merger masses, SIFT measures, and  $\Pi_N$  variables all important to QCD vs Dark QCD discrimination power
- ❑ Both late and early merger variables are important to the analysis
- ❑ AUROC score of 83% using only these SIFTy variables

Receiver Operating Characteristic for Validation Fold 1  
Background vs. Signal



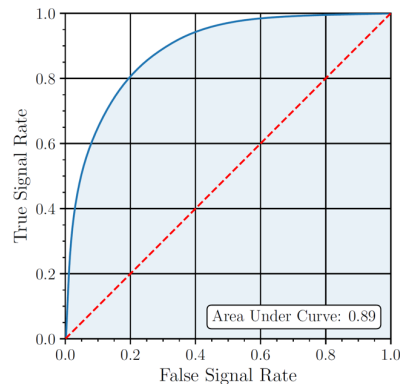
Feature Importance to Total Gain in Training Fold 1  
Background vs. Signal



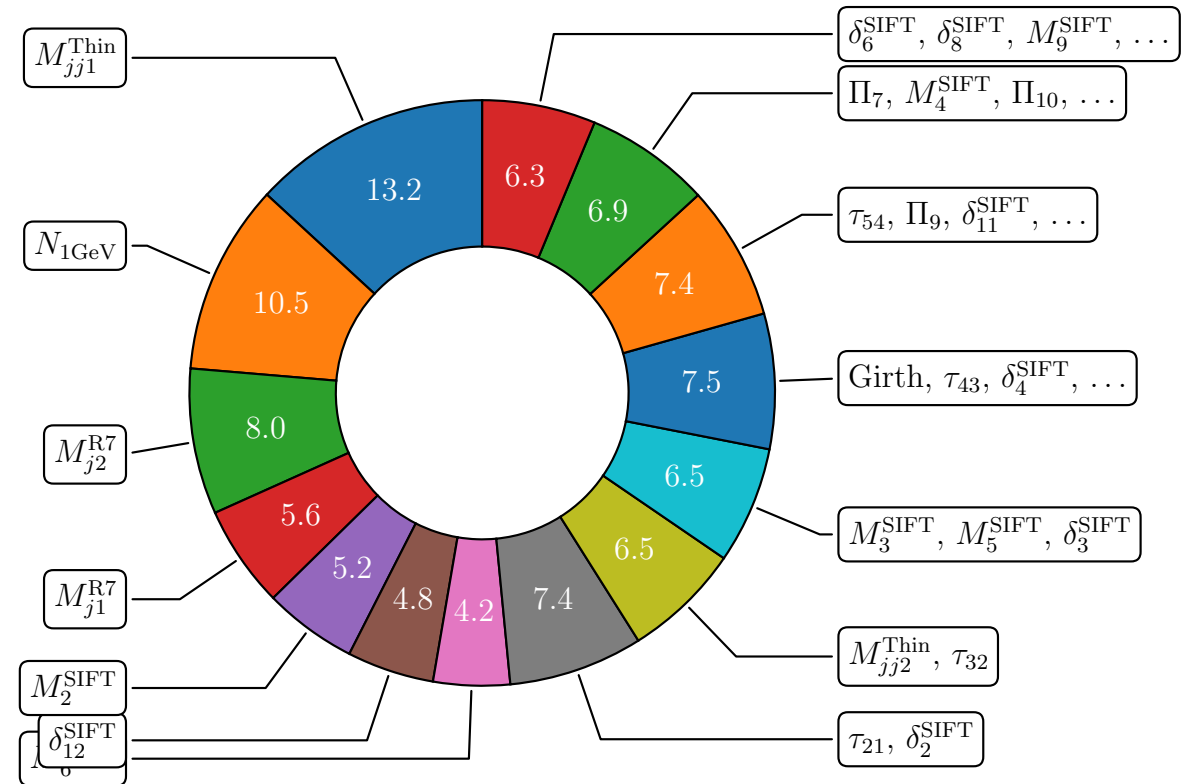
# BDT Results $M_{\pi} = 10 \text{ GeV}$ – All Variables

- ❑ A kitchen-sink analysis utilizes dijet masses, monojet masses, particle counts, and SIFTy variables
- ❑ Uses information very similar to ATLAS analysis, improved by addition of the two other approaches
- ❑ AUROC score of 89%

Receiver Operating Characteristic for Validation Fold 1  
Background vs. Signal



Feature Importance to Total Gain in Training Fold 1  
Background vs. Signal



# Classifier scores

$M_\pi$	Strassler All	SIFT All	Classic QCD	Kitchen Sink
10	79	83	77	89
25	89	94	87	96
50	95	98	91	99
120	98	99	93	100
200	99	99	92	100
500	96	99	73	99

- ❑ SIFTy technique alone is outcompeting dijet resonance techniques of Strassler and broadly-classified QCD variables alone
- ❑ Putting it all together, we have strong classification power throughout the explored parameter space
  - Makes explicit the complementarity of these approaches



# Conclusion

- ❑ We can achieve this level of discrimination between QCD showers and new dark QCD showers with prompt decays
- ❑ We can do that in this regime, where it seems other techniques do not work as well
  - Information from these techniques is complimentary
- ❑ This is a valuable expansion of the reach of the LHC into the hidden valley parameter space

# Exploring Further

- Explore using this as a potential discovery tool rather than a discrimination tool
- Broaden our parameter space away from the fixed  $Z'$  mass
- Cross train machine learning against each individual mass

Thank You

