Identifying the Quantum Color Representation of New Particles with Machine Learning

John Kruper[1]

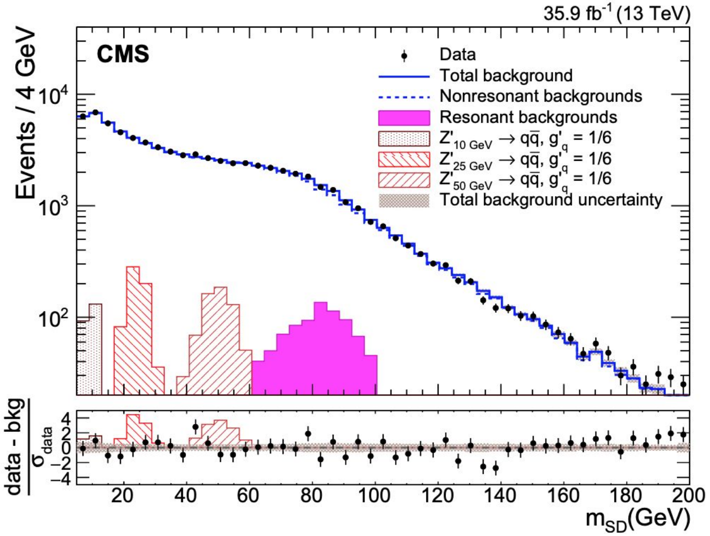
With Jakub Filipek[1], Shih-Chieh Hsu[1], Kirtimaan Mohan[2], and Benjamin Nachman[3]

[1] University of Washington, [2] Michigan State University, [3] Lawrence Berkeley National Laboratory

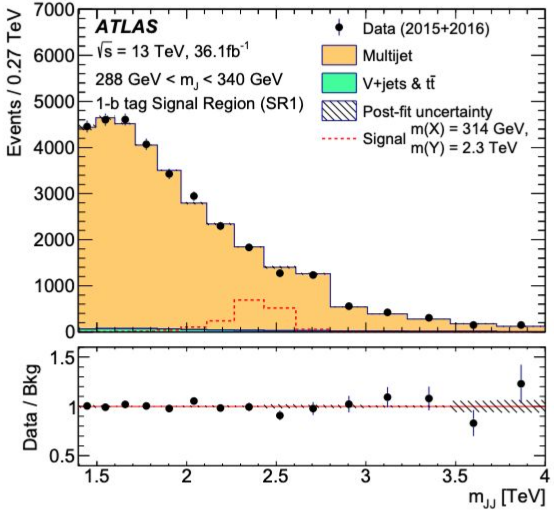
2019 APS Division of Particles & Fields Meeting
Northeastern University, July 29, 2019

New Particles Search in Hadronic Decay

- Many theories predict new particles that decay hadronically



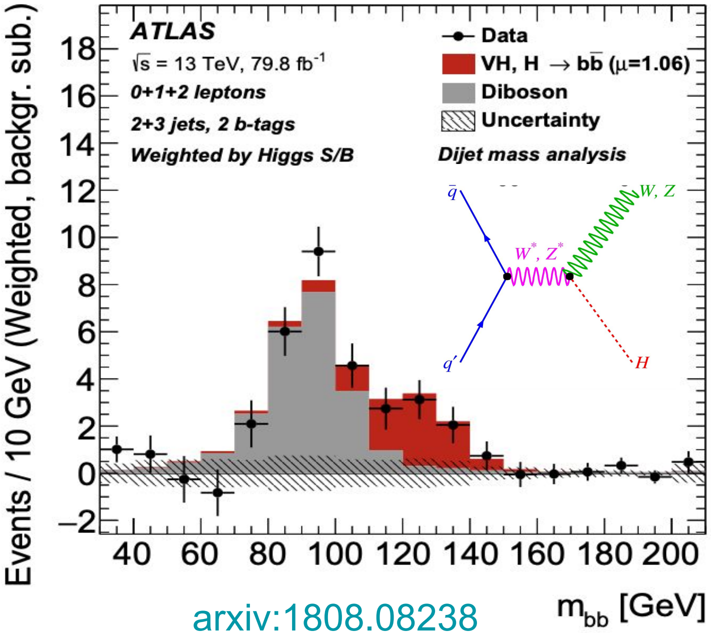
arxiv:1905.10331



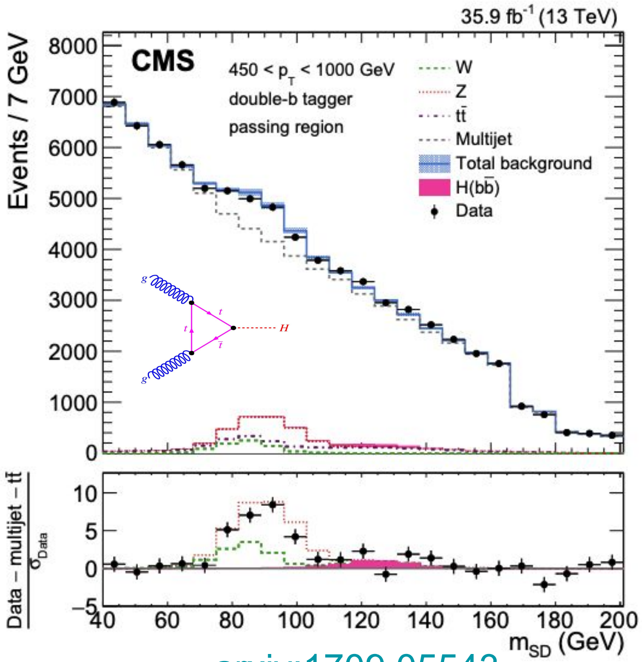
arxiv:1709.06783

- If discovered, how can we learn about the properties of these new particles?

Beyond Standard Model (BSM) Particles at 125 GeV



arxiv:1808.08238



arxiv:1709.05543

- Higgs was discovered in 2012
- As luminosity increases, what if there are excess events at the 125 GeV peak?
- How can we distinguish the Higgs from other new particles at the same mass?

BSM Models

M_j = 125 GeV for all models

Resonance	Interaction	J	$SU(3)_C$	$\|Q_e\|$	Decay
Higgs Boson H	$g_{Hff}\bar{f}fH$	0	0	0	qq~
Higgs Boson H	$\frac{1}{v}g_{ggH}\ H\,G^{\mu\nu}G_{\mu\nu}$	0	0	0	gg
Leptophobic Z'	$\frac{g_B}{6}\bar{q}\gamma^\mu qZ'_\mu$	1	0	0	qq~
Coloron C_μ	$g_s\tan\theta\ \bar{q}\ T^a\gamma^\mu qC^a_\mu$	1	8	0	qq~
Excited quark q^*	$\frac{1}{2\Lambda}\bar{q}^*_R\sigma^{\mu\nu}[g_S f_S\frac{\lambda^a}{2}G^a_{\mu\nu}]q_L$	1/2	3	2/3	qg
Octet Scalar S_8	$\frac{g_s d_{ABC}k_s}{\Lambda}S^A_8 G^B_{\mu\nu}G^{C,\mu\nu}$	1	8	0	gg

arxiv:1710.04661

Event Generation

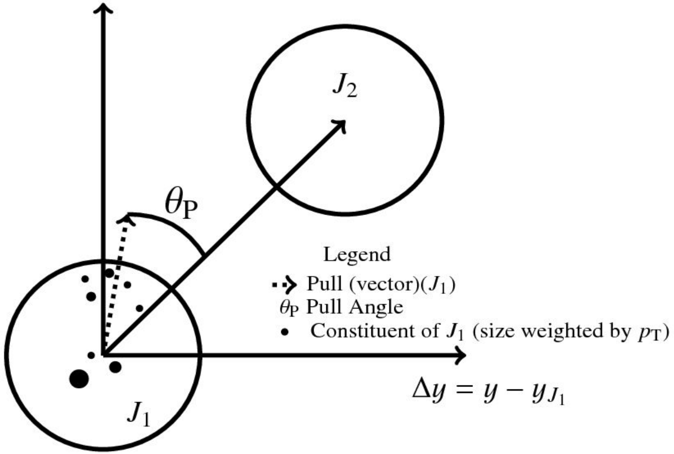
- MadGraph5 version 2.5.5 and Pythia8 version 8.235
    - $pp \rightarrow X (\rightarrow jj)\, j$; $X \in \{H, Z', C_\mu, q^*, S_8\}$
- FastJet 3 version 3.3.2
    - $R = 1.0$
    - Trimmed with $R = 0.3$ subjets and $p_T^{subjet} < 0.05 \times p_T^{jet}$
- Jet Selection
    - $|\eta|<2.0$
    - $300\ \text{GeV} < p_T^{jet} < 600\ \text{GeV}$
    - $100\ \text{GeV} < m^{jet} < 150\ \text{GeV}$

Similar to generation in: arxiv:1710.04661

High Level Tagger: Jet Pull Definition

Jet pull is a pT-weighted radial moment:

$$\sum_{i \in J} \frac{p_T^i |\vec{r_i}|}{p_T^J} \vec{r_i}$$



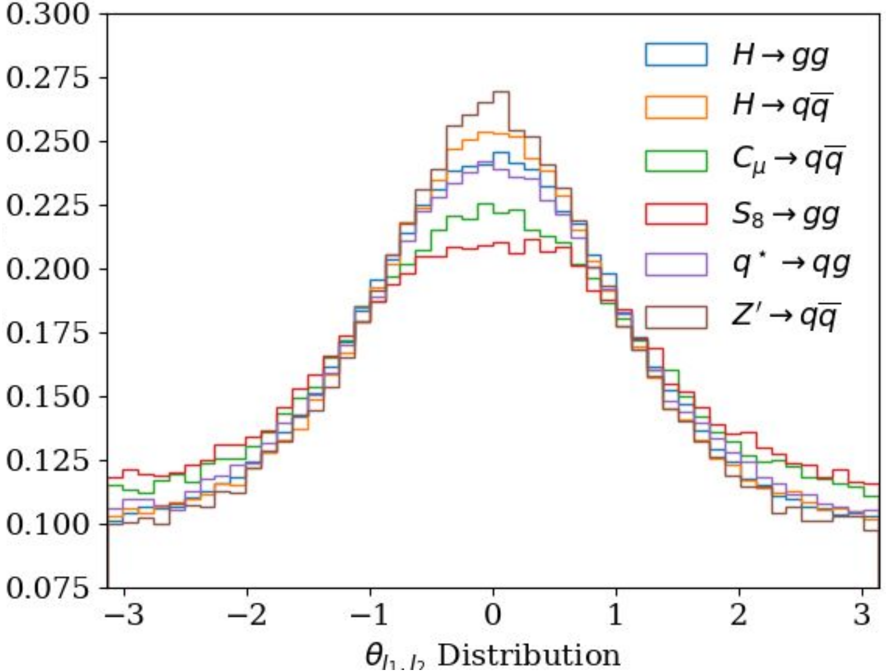
TOPQ-2014-09/

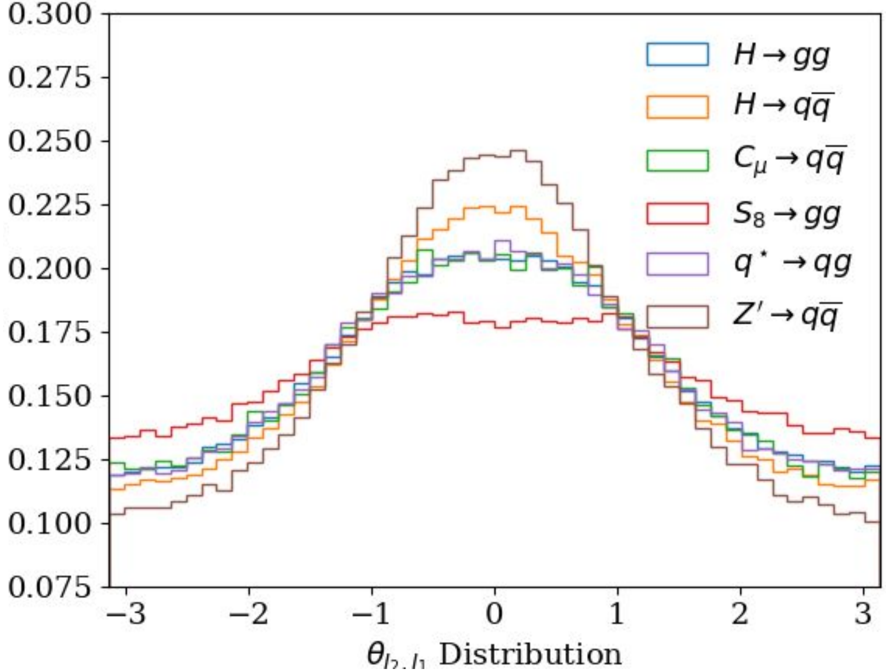
arxiv:1001.5027

Jet pull is the state-of-the-art variable motivated by QCD and a classical choice to capture color information

Jet Pull Histograms

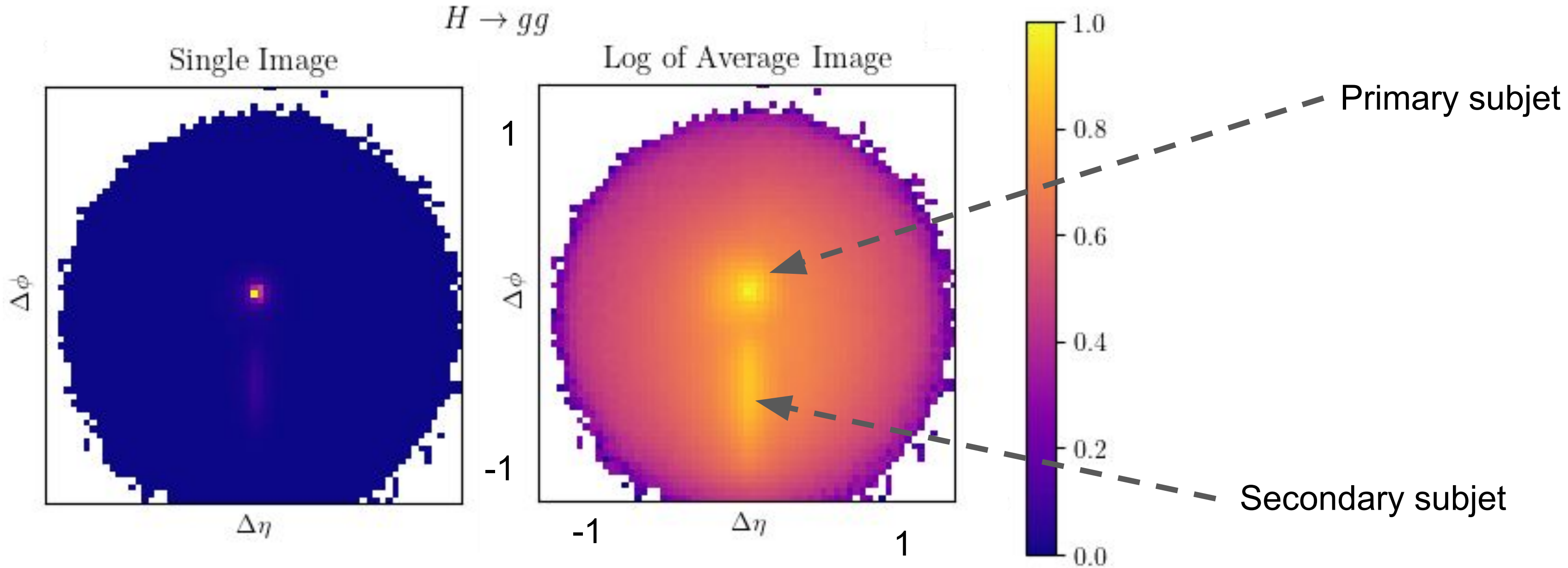
Larger color numbers (ie: S_8, C_μ) are more spread out!



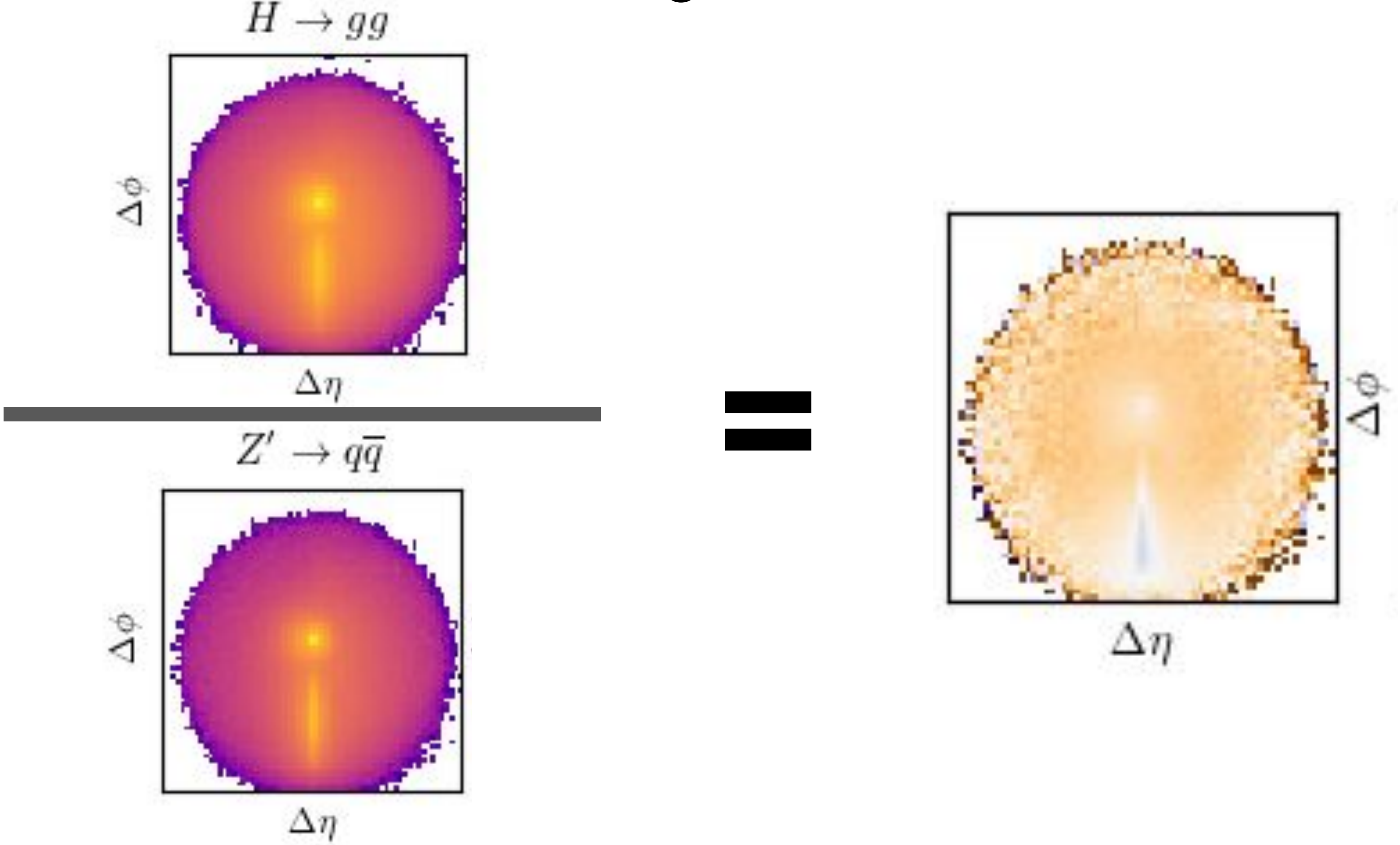


Jet-Images

- Jet-Images: Energy depositions of the particles are the pixel intensities
- Images are centered, pixelated to 65x65 pixels, rotated, logged, and normalized:



Jet-Images, Ratio



Jet-Images, All Ratios

Red Region has only quark-antiquark decay and are similar to each other

Blue box has gluon in decay and similar jet-images to H→gg

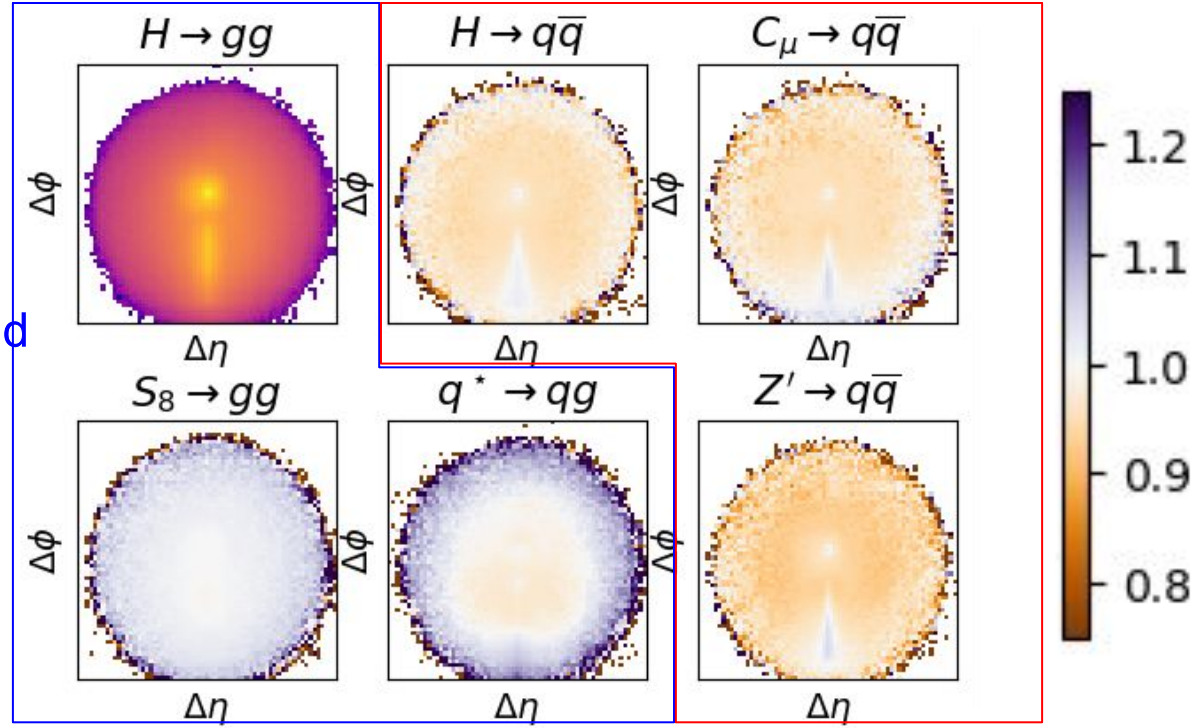


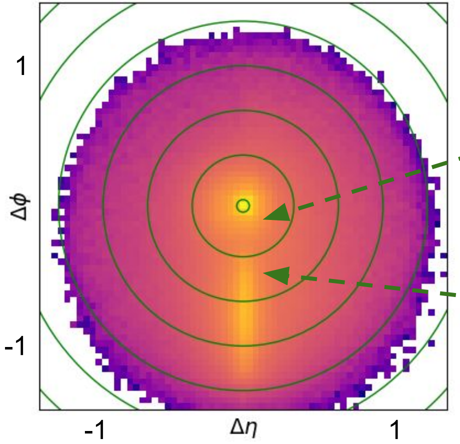
Image of H->gg and Ratios to H->gg

High Level Tagger: Energy Flow (E-flow)

Idea: Sum pixel energies in rings of the jet-image to capture pattern

- The first circle has R=0.015, the rings increase R by 0.1
- The rings we chose are shown on the picture below

$$E_n = \sum_{ij} E \text{ for all } i, j$$
$$\text{where } R_{n-1} < \sqrt{i^2 + j^2} \leq R_n$$



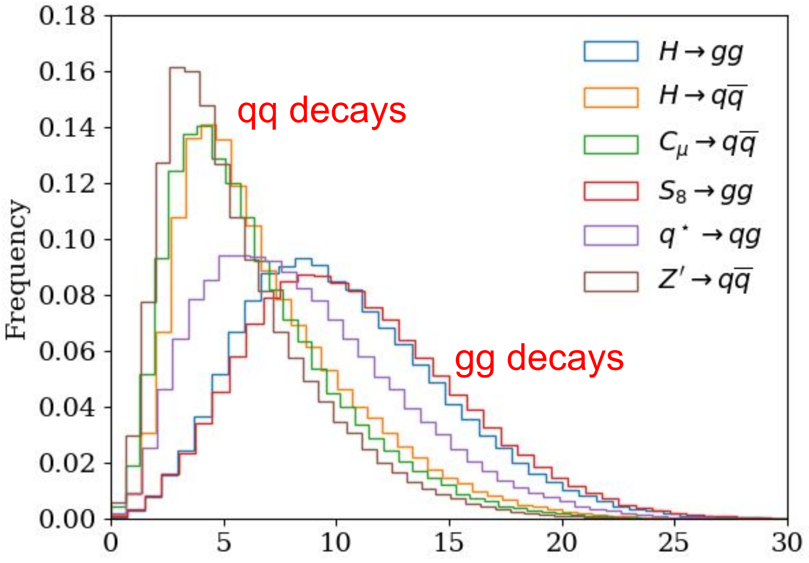
E_2 is a part of the primary trimmed subjet

E_3 is a region between the primary and secondary subjet

E-flow Histograms

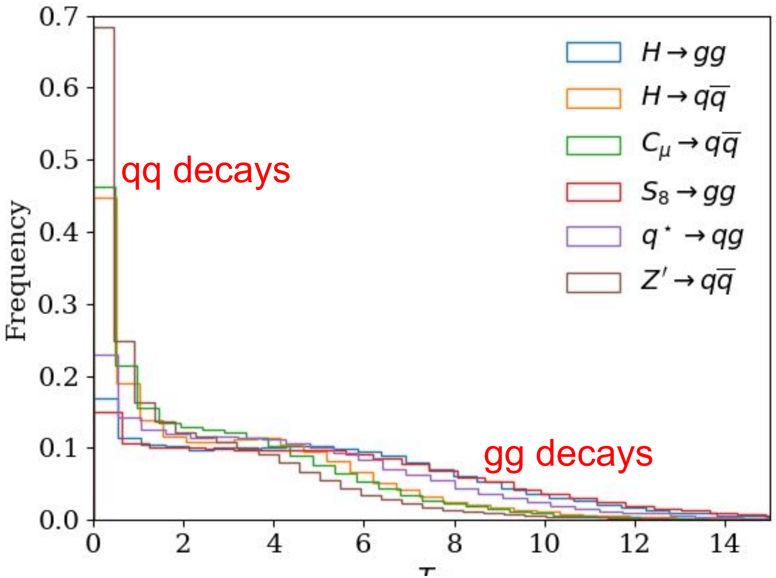
- Particles with gluons in the decay leave more deposits in the second and third rings (E_2 and E_3)

E_2 (Around Primary subjet)



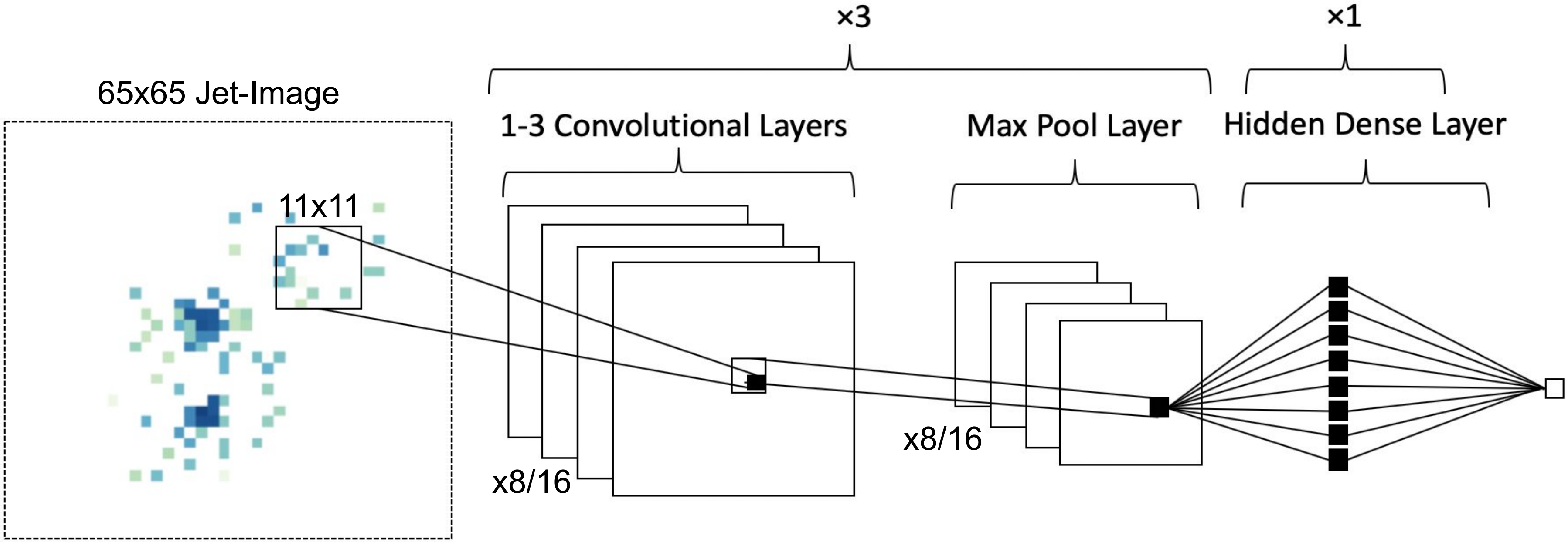
Normalized sum of energies

E_3 (Between subjets)



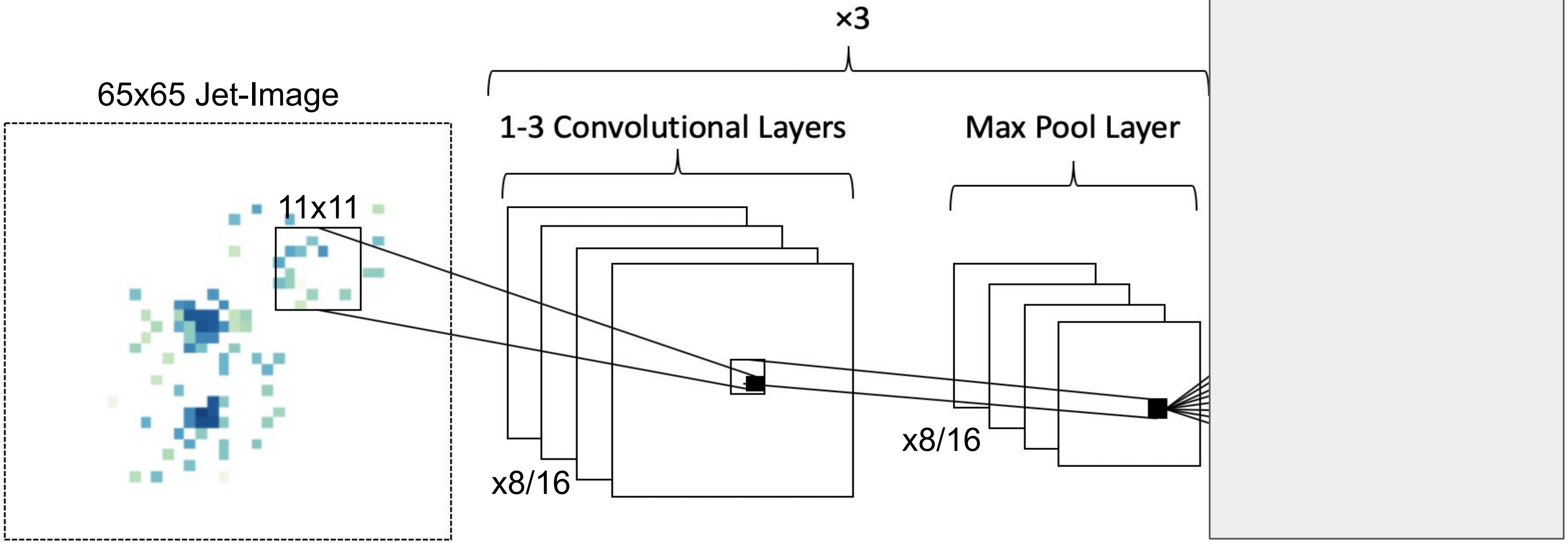
Normalized sum of energies

Low Level Tagger: Convolutional Neural Network (CNN)



- First layer has 8 11×11 filters, others have 16 3×3 filters
- Layer 3, 5, and 7 followed by 2×2 max-pooling layers
- Dense Layer has width 128 followed by 0.5 dropout

CNN Optimization, Choices



- Used Sequential Model-based Algorithm Configuration (SMAC3) to optimize hyperparameters
- Large filter in first layer + many convolutional layers
  - learns relations between the sparse, non-zero parts of image

CNN Optimization, Choices

- Single, wide fully connected layer
  - Final decisions are not complicated

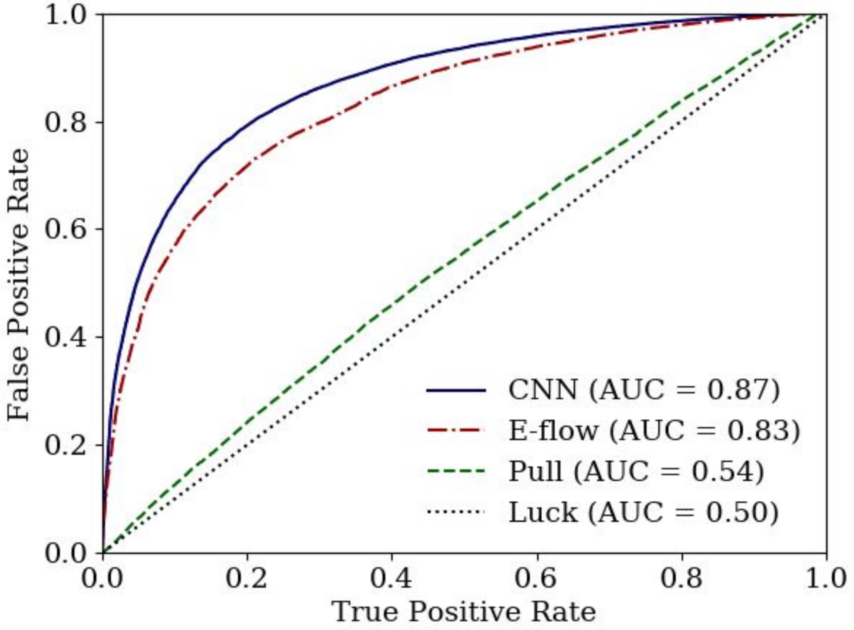
Training

- For each of the 6 processes, we have 150,000 events
- We create 15 combinations of 300,000 events
- One process in each combination is chosen as signal
  - Signal mass is reweighted to be the same distribution as the background mass
- For jet pull and e-flow, we use a boosted decision tree and a 80-20% train-test split
  - We used Adaboost from scikit to implement the boosted decision tree
- For CNN, we do another training-validation-test split of 64-16-20%
  - Made with Keras built on Tensorflow

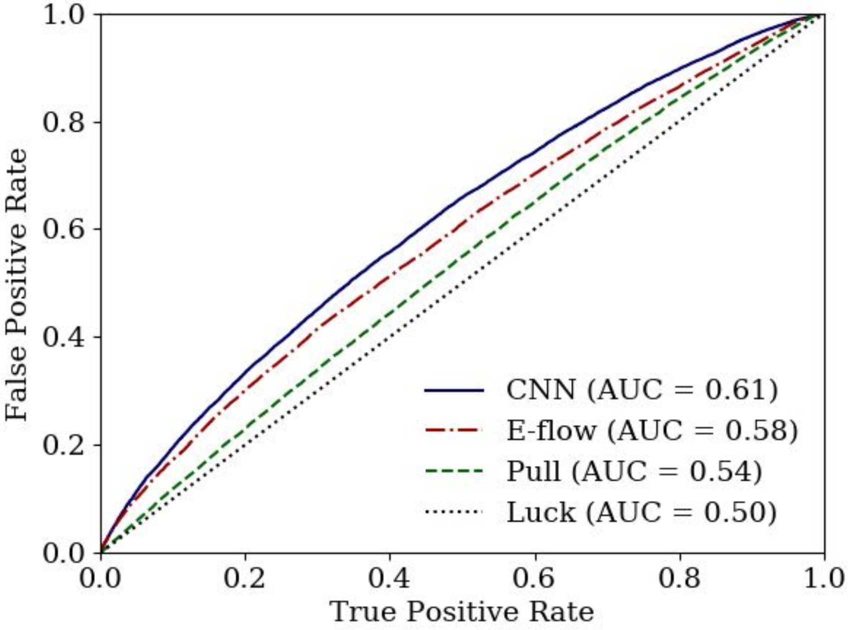
Training results, ROC curves

$$\text{True Positive Rate} = \frac{\text{True Positives}}{\text{All Positives}}, \quad \text{False Positive Rate} = \frac{\text{False Positives}}{\text{All Positives}}$$

Z’→qq (True), H→gg (False)



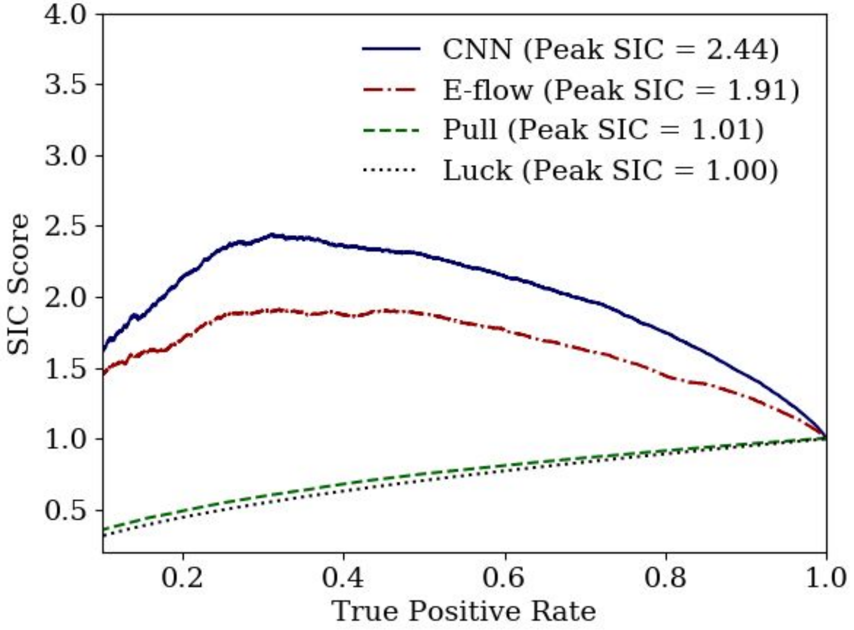
S_8→gg (True), H→gg (False)



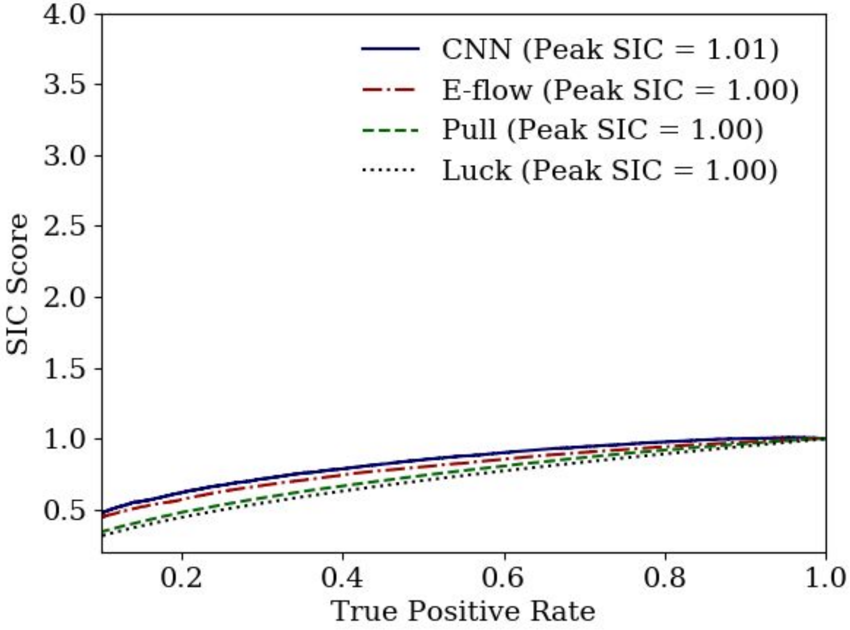
Significance improvement characteristic (SIC) Curves

$$\text{SIC Score: } \frac{\text{True Positive Rate}}{\sqrt{\text{False Positive Rate}}}$$

Z'→qq (signal), H→gg



S_8→gg (signal), H→gg



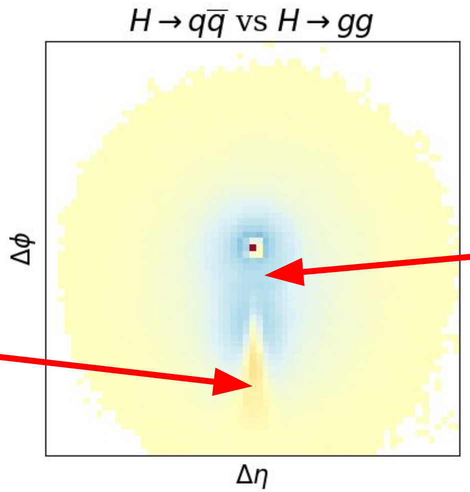
Pearson Correlation Coefficient Images

For each pixel, calculate:

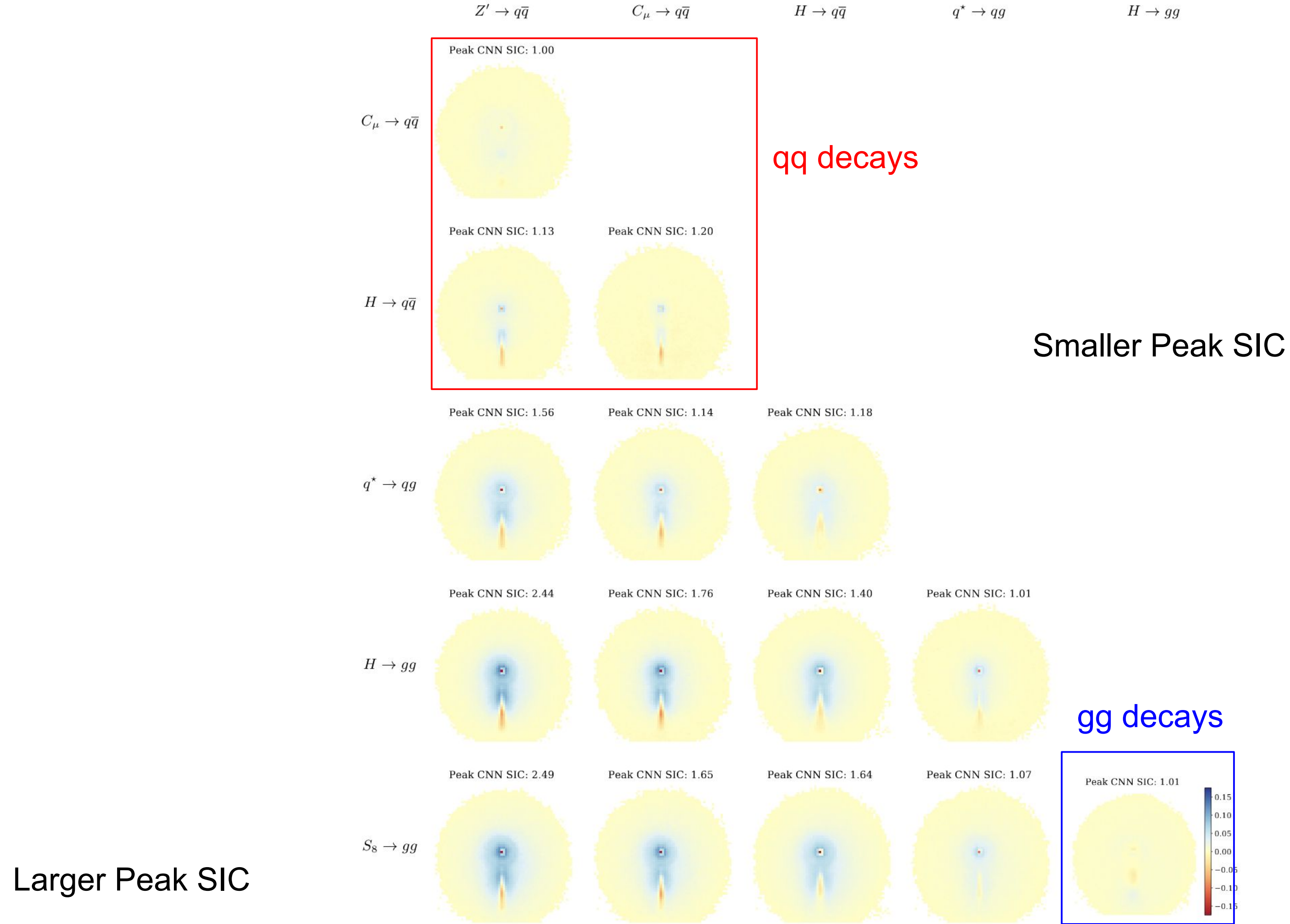
$$\rho_{X,Y} = \frac{\mathrm{E}[(X - \mu_X)(Y - \mu_Y)]}{\sigma_X \sigma_Y}$$

where X is that pixel's distribution and Y is the distribution of true label

Pixel activations in the red regions linearly correlate with being a H->qq process

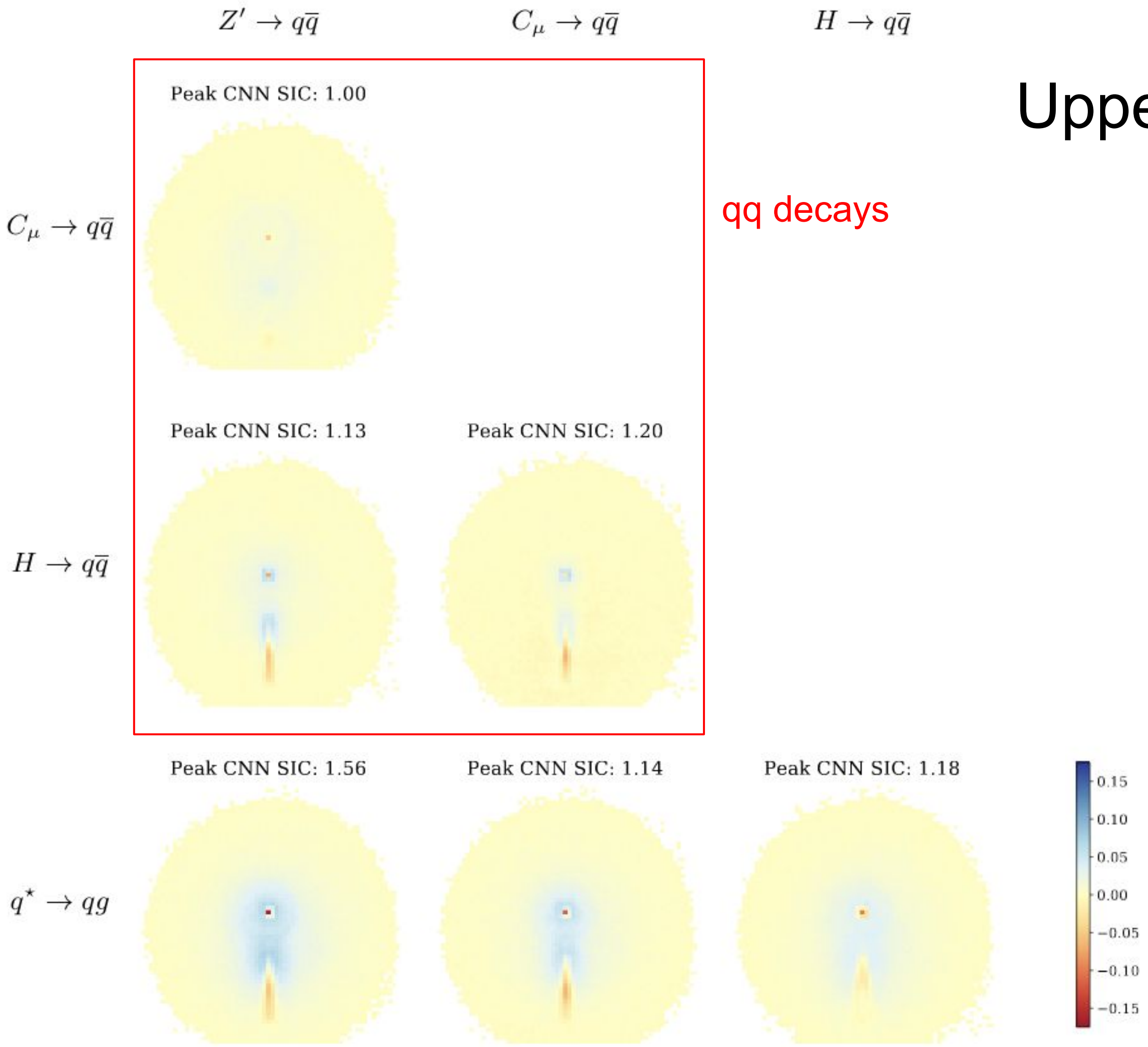


Pixel activations in the blue regions linearly correlate with being a H->gg process



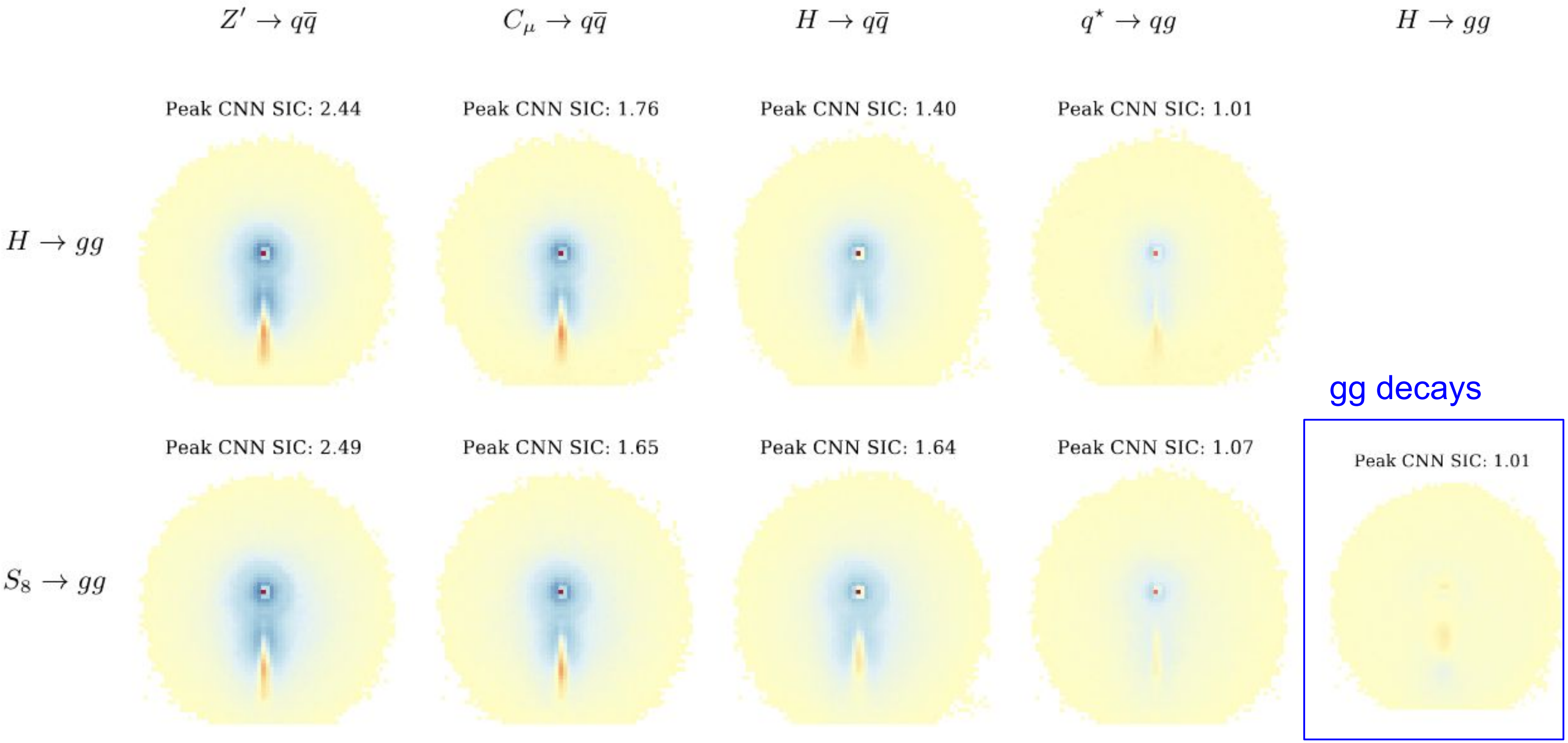
$Z' \to q\bar{q}$
$C_\mu \to q\bar{q}$
$H \to q\bar{q}$
$q^\star \to qg$
$H \to gg$
$C_\mu \to q\bar{q}$
$H \to q\bar{q}$
$q^\star \to qg$
$H \to gg$
$S_8 \to gg$
Peak CNN SIC: 1.00
Peak CNN SIC: 1.13
Peak CNN SIC: 1.20
Peak CNN SIC: 1.56
Peak CNN SIC: 1.14
Peak CNN SIC: 1.18
Peak CNN SIC: 2.44
Peak CNN SIC: 1.76
Peak CNN SIC: 1.40
Peak CNN SIC: 1.01
Peak CNN SIC: 2.49
Peak CNN SIC: 1.65
Peak CNN SIC: 1.64
Peak CNN SIC: 1.07
Peak CNN SIC: 1.01
qq decays
gg decays
Smaller Peak SIC
Larger Peak SIC

Upper Half



$Z' \to q\bar{q}$
$C_\mu \to q\bar{q}$
$H \to q\bar{q}$
$C_\mu \to q\bar{q}$
$H \to q\bar{q}$
$q^\star \to qg$
Peak CNN SIC: 1.00
Peak CNN SIC: 1.13
Peak CNN SIC: 1.20
Peak CNN SIC: 1.56
Peak CNN SIC: 1.14
Peak CNN SIC: 1.18
qq decays
0.15
0.10
0.05
0.00
−0.05
−0.10
−0.15

Lower Half

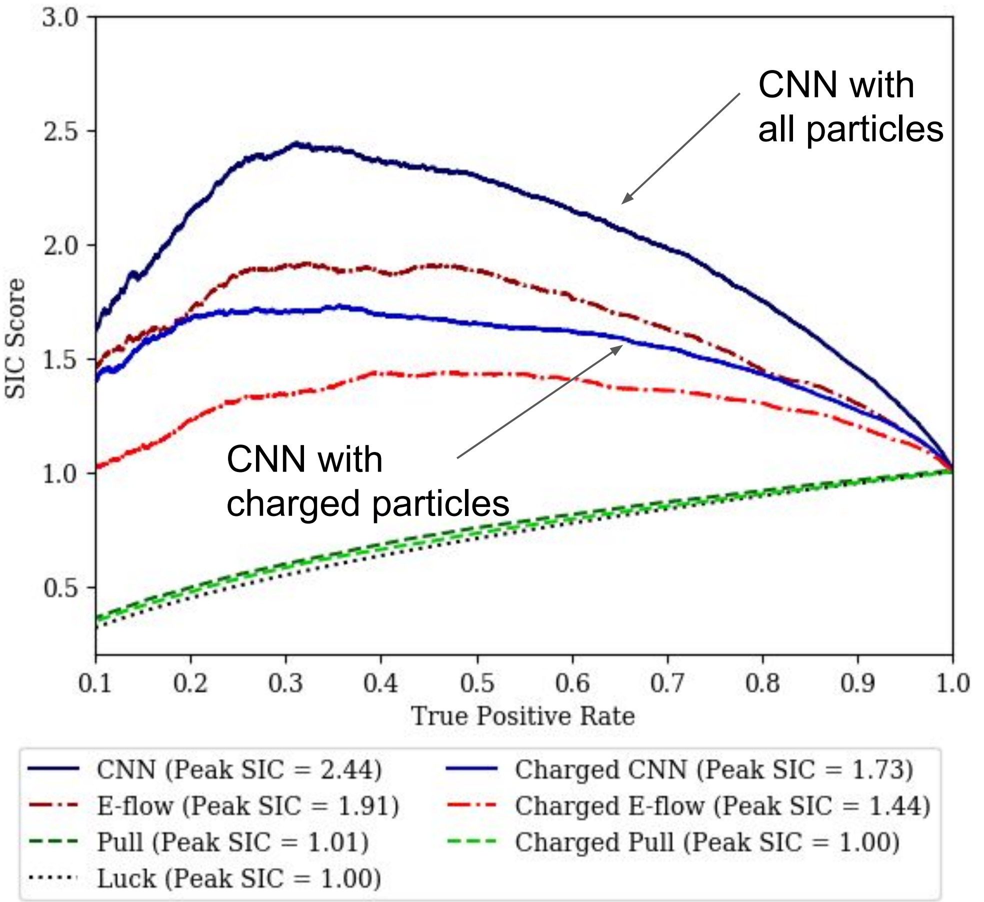


$Z' \to q\bar{q}$
$C_\mu \to q\bar{q}$
$H \to q\bar{q}$
$q^\star \to qg$
$H \to gg$
$H \to gg$
$S_8 \to gg$
Peak CNN SIC: 2.44
Peak CNN SIC: 1.76
Peak CNN SIC: 1.40
Peak CNN SIC: 1.01
Peak CNN SIC: 2.49
Peak CNN SIC: 1.65
Peak CNN SIC: 1.64
Peak CNN SIC: 1.07
gg decays
Peak CNN SIC: 1.01

Dependences

- Variations on pythia settings have no impact on performance

- When using only charged particles, there is a decrease in performance (shown right)



Conclusion

CNN trained on jet-images is a powerful method to distinguish particles with different radiation patterns

- Significantly better than high-level jet-pull tagger
- More specifically, particles with final states qq~ are difficult to distinguish
  - For example: H, C_μ, and Z’
- Particles with final states with g in them are difficult to distinguish
  - For example, Hadronic H, S_8, q^*

Acknowledgement

John Kruper, Jakub Filipek, and Shih-Chieh Hsu were supported by the Department of Energy, Office of High Energy Physics Early Career Research program under Award Number DE-SC0015971

Benjamin Nachman was supported by the U.S. Department of Energy, Office of Science under contract DE-AC02-05CH11231.

Identifying the Quantum Color Representation of New Particles with Machine Learning

Jakub Filipek,[1,*] Shih-Chieh Hsu,[1,†] John Kruper,[1,‡] Kirtimaan Mohan,[2,§] and Benjamin Nachman[3,¶]

[1]*Department of Physics, University of Washington, Seattle, WA 98195, USA*
[2]*Department of Physics and Astronomy Division,*
Michigan State University, 567 Wilson Rd, East Lansing, MI 48824, USA
[3]*Physics Division, Lawrence Berkeley National Laboratory, Berkeley, CA 94720, USA*

With the great promise of deep learning, discoveries of new particles at the Large Hadron Collider (LHC) may be imminent. Following the discovery of a new particle in an all-hadronic channel, deep learning can also be used to identify the quantum numbers of the new particle. Convolutional neural networks (CNNs) using jet-images can significantly improve upon existing techniques to identify the quantum chromodynamic (QCD) representation ('color') of a two-prong jet using its substructure. Additionally, jet-images are useful in determining what information in the jet radiation pattern is useful for classification, which could inspire future taggers. These techniques improve the categorization of new particles and are an important addition to the growing jet substructure toolkit, for searches and measurements at the LHC now and in the future.

Preprocessing Steps

- Preprocessing:
  - Translate so leading subjet is origin
  - Pixelate to 65x65 pixels
  - Rotate so second subjet is directly underneath the origin
  - Flip horizontally so one side always has the most energy
  - Log and normalize

Peak SIC for all Combinations and Models

	Model	$H \to gg$	$H \to q\overline{q}$	$C_\mu \to q\overline{q}$	$S_8 \to gg$	$q^\star \to qg$	$Z\prime \to q\overline{q}$
$H \to gg$	CNN		**1.4041**	**1.7571**	**1.0103**	**1.0108**	**2.4413**
	T-jets		1.2381	1.4955	1.0043	1.0053	1.9128
	Pull		1.0000	1.0049	1.0041	1.0041	1.0050
$H \to q\overline{q}$	CNN	**1.4041**		**1.2049**	**1.6372**	**1.1791**	**1.1313**
	T-jets	1.2381		1.0081	1.3710	1.0575	1.0131
	Pull	1.0000		1.0043	1.0041	1.0044	1.0048
$C_\mu \to q\overline{q}$	CNN	**1.7571**	**1.2049**		**1.6471**	**1.1405**	**1.0000**
	T-jets	1.4955	1.0081		1.4562	1.0532	1.0034
	Pull	1.0049	1.0043		1.0000	1.0000	1.0000
$S_8 \to gg$	CNN	**1.0103**	**1.6372**	**1.6471**		**1.0663**	**2.4969**
	T-jets	1.0043	1.3710	1.4562		1.0515	1.9647
	Pull	1.0041	1.0041	1.0000		1.0000	1.0000
$q^\star \to qg$	CNN	**1.0108**	**1.1791**	**1.1405**	**1.0663**		**1.5619**
	T-jets	1.0053	1.0575	1.0532	1.0515		1.2381
	Pull	1.0041	1.0044	1.0000	1.0000		1.0003
$Z\prime \to q\overline{q}$	CNN	**2.4413**	**1.1313**	**1.0000**	**2.4969**	**1.5619**	
	T-jets	1.9128	1.9128	1.0034	1.9647	1.2381	
	Pull	1.0050	1.0050	1.0000	1.0000	1.0003	