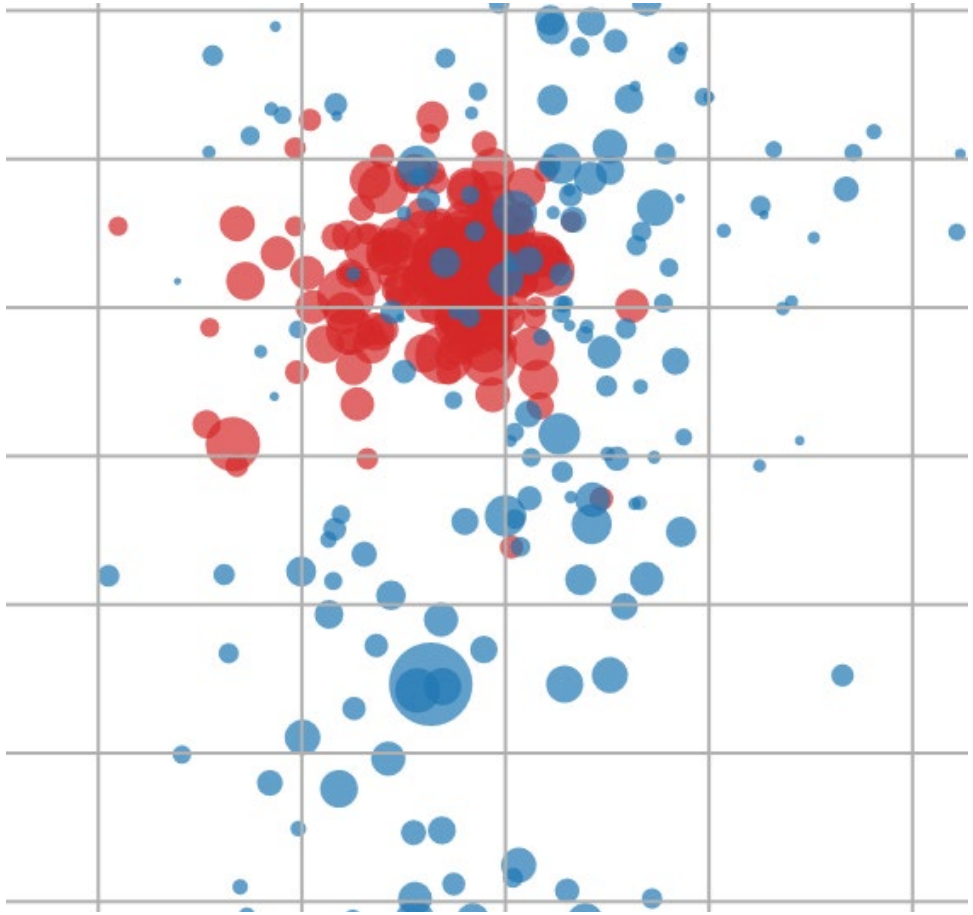
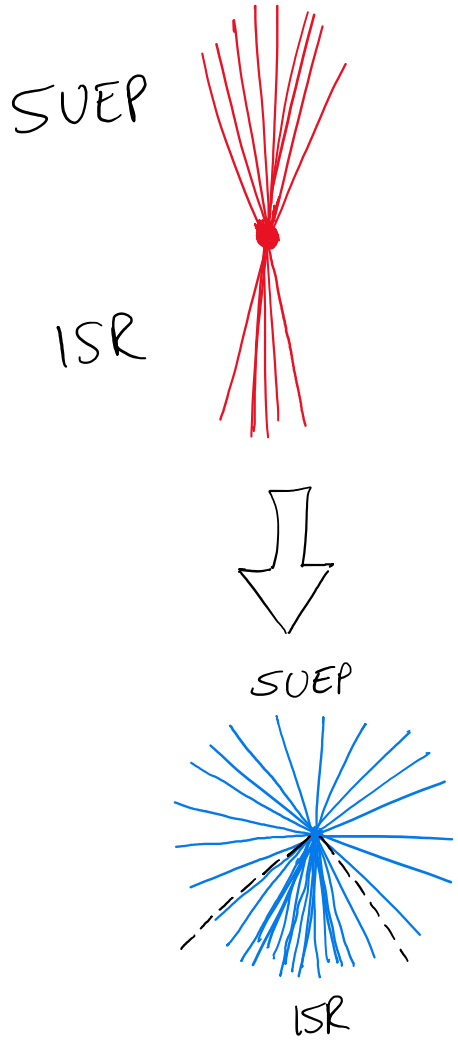


ML for SUEP Analysis

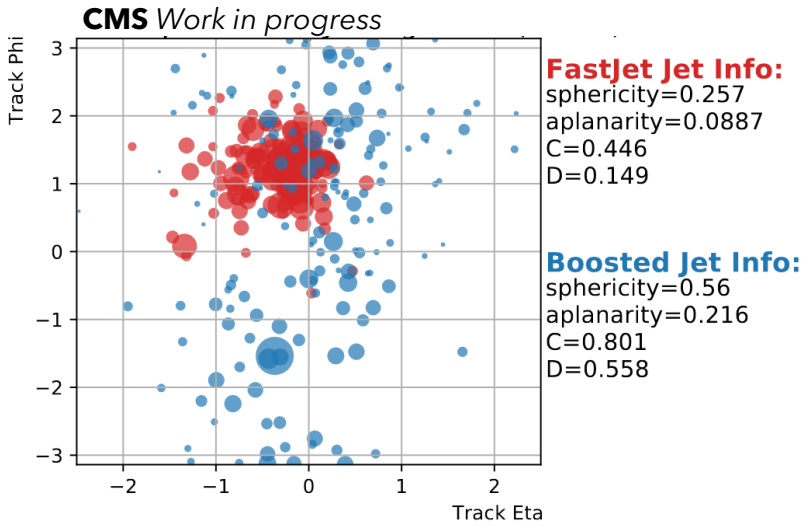
Luca Lavezzo for the CMS Collaboration

SUEPs

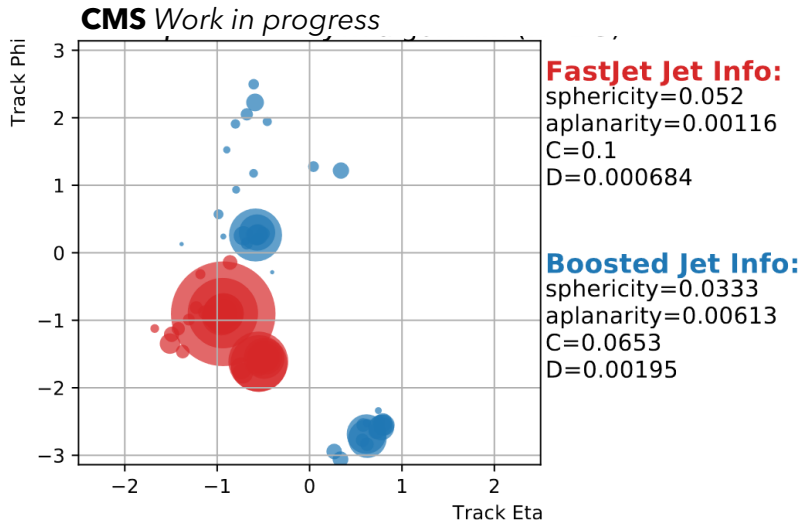
- Large t'Hooft couplings → high multiplicity, spherically symmetric jets (SUEPs)
- Can look like soft pile up
- For most mediator masses, can be targeted with **number of tracks** and **sphericity** in frame of the boosted jet (boost needed due to trigger)



SUEP



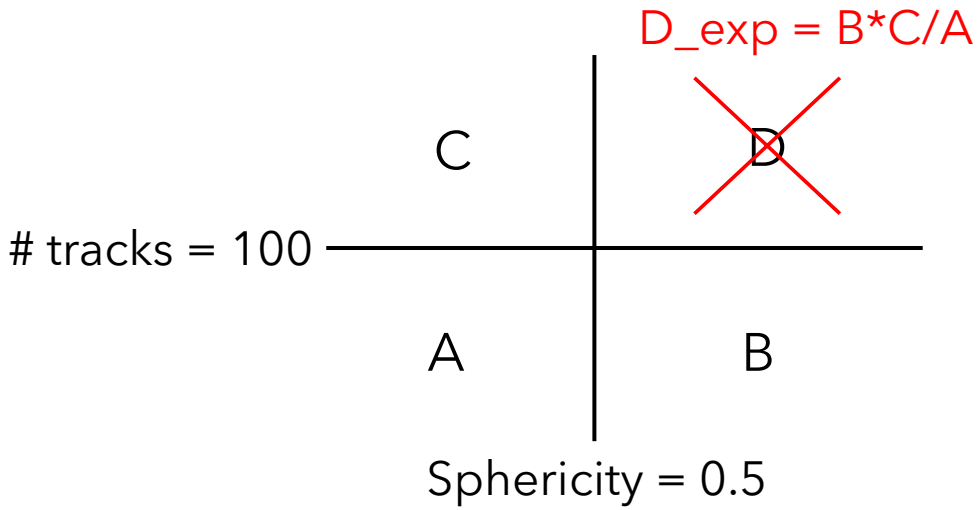
QCD MC

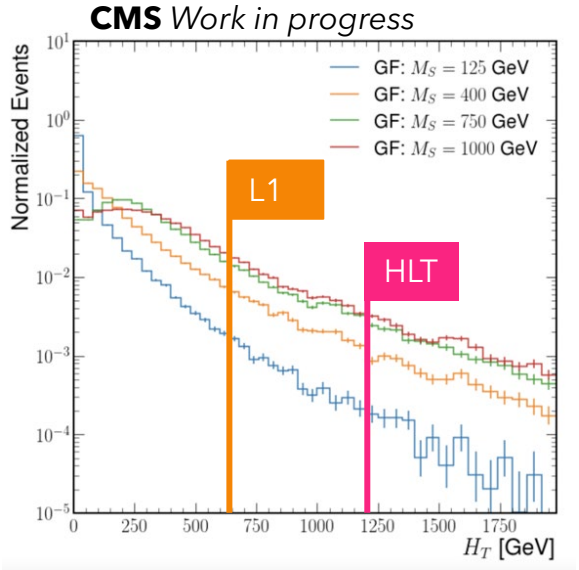


Cut-Based Analysis

ABCD Method

+ Uncorrelated variables to predict background in signal region using data
+ Discriminators: (1) sphericity of SUEP tracks (2) # tracks in event
  + Many more variables that might help the classification task





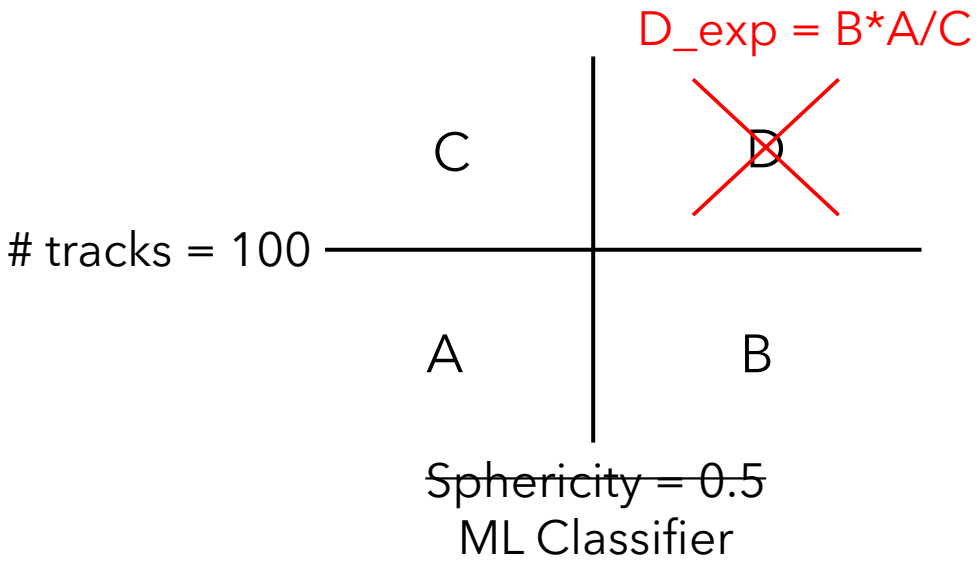
H_T calculated from gen-level jets ($p_T > 20\ GeV, |\eta| < 4.7$) for different SUEP mediators through Gluon Fusion (GF) production mechanism.

+ Already low efficiency for signal acceptance even from L1 trigger for low mass mediators
+ Looking to enhance sensitivity for low mass mediators where distributions are in bulk of QCD
+ Try ML!

  Couple options for classification task: **CNNs, GNNs**

DiSco [arXiv:2007.14400v1]

- Even using ML, want data-driven background estimation with ABCD, cannot rely on QCD in this tail-end of the phase space: DiSco employed
- Replace engineered feature(s) with deep-learning constructed discriminator(s)
  - Single DiSco: 1 engineered feature, 1 ML discriminator
  - Double DiSco: 2 ML discriminators



- Decorrelation between variables enforced at the loss function:

$$\mathcal{L}[f(X)] = \mathcal{L}_{\text{classifier}}[f(X), y] + \lambda \, \text{dCorr}^2_{y=0}[f(X), X_0]$$

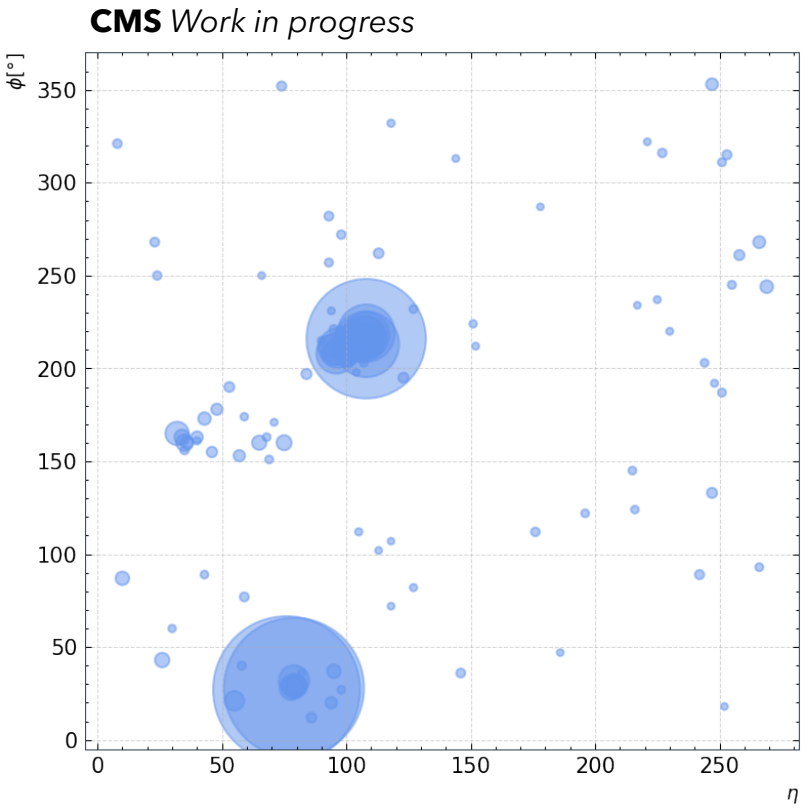
From: [arXiv:2007.14400v1]

ResNet

+ Model `torchvision.models.resnet18` used with transfer learning weights `ResNet18_Weights.IMAGENET1K_V1`
+ **Single DiSco** using # tracks as second ABCD variable

+ Tracker (ϕ, η) info of PFCands as image of size (360, 280)
+ pT of tracks used to normalize the pixel values

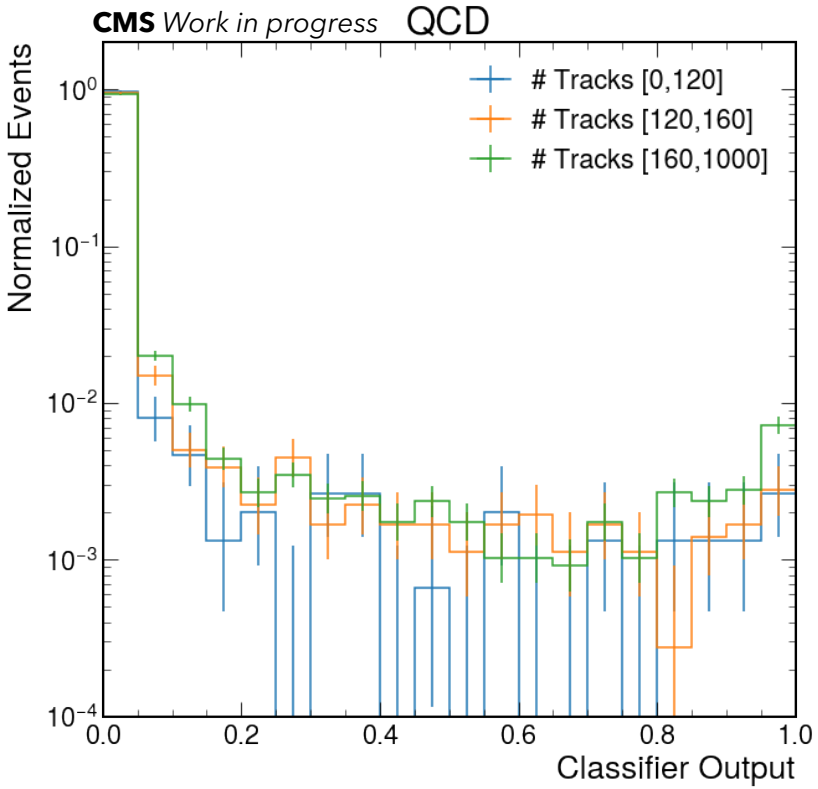
+ Training and validation data: SUEP m125 (darkPhoHad), QCD*
+ Event selection: HT of ak8jets > 600GeV to match scouting data
+ Track selection: >1 GeV, $|\eta| < 2.4$
+ Output labels: [0,1] class prediction per image



Sample MC QCD event used in training. Tracks are shown by disks scaled by their pTs.

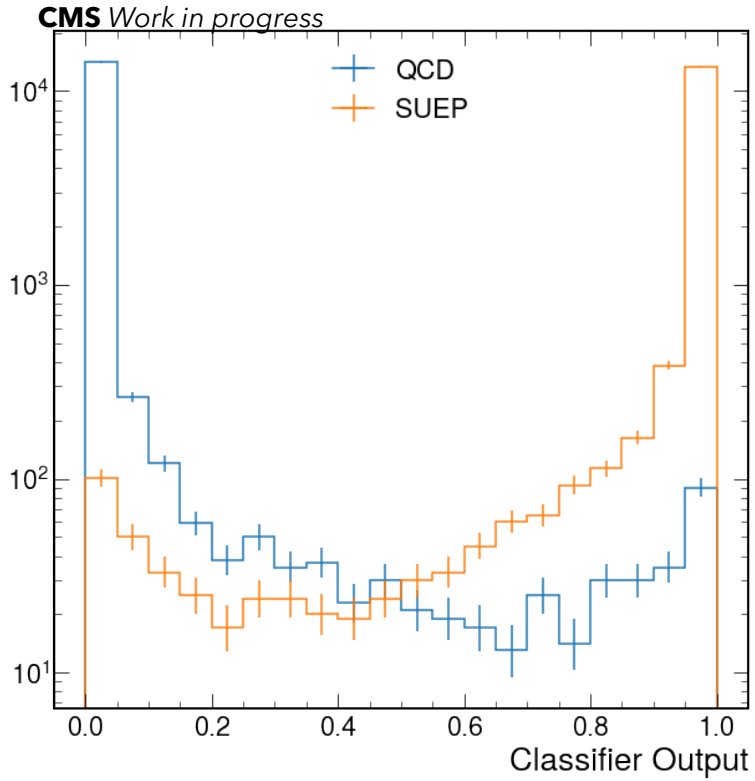
*QCD = QCD_Pt-15to7000_TuneCP5_Flat2018

ResNet



1D slices in # tracks as function of classifier output shows the DiSco decorrelation is working, as the slices have the same functional shape

Classification is very effective for the lowest mass (125 GeV), which was problematic in cut-based analysis, with SUEPs consistently predicted at 1 and QCD as 0



*Both plots show model performance on previously unseen validation data.

Summary

- ResNet + DiSco
  - Effective for classification + data-driven background estimation

Future steps:
- Train for HT > 1200 GeV data and run inference
- Scale up the inference to full HT > 600 GeV samples
- Double DiSco

- Other models:

JetSSD: had worse performance than ResNet in classification of events.

ParticleNet: in the works!

Thank you to the whole SUEP group!

Christoph Paus, Chad Freer, Luca Lavezzo (MIT), Benedikt Maier, Maurizio Pierini (CERN), Adrian Pol, Isobel Ojalvo (Princeton), Carlos Erice Cid, Zeynep Demiragli (BU), Sarah Eno, Christos Papageorgakis (UMD), Tova Holmes, Benjamin Thornberry (University of Tennessee Knoxville), Karri DiPetrillo, Kevin Pedro (Femilab), Tres Reid (Cornell)

Backup

DiSco Loss [arXiv:2007.14400v1]

- Decorrelation between variables enforced:

$$\mathcal{L}[f(X)] = \mathcal{L}_{\text{classifier}}[f(X), y] + \lambda \, \text{dCorr}^2_{y=0}[f(X), X_0]$$

From: [arXiv:2007.14400v1]

- *f*(*X*) model *f* acting on variables *X*
- *y* as truth labels (0 = QCD, 1 = SUEP)
- λ scaling hyperparameter, needs to be tuned
- *dCorr^2* correlation function
  - Acts only on y=0 (QCD), since we only care to decorrelate background data
  - Measure of dependence between random vectors
  - dCorr^2 = dCov^2(f, g) / dCov(f, f) * dCov(g, g)
  - dCov = distance covariance