The Role of Jet Geometry in Jet Classification

Understanding and Improving jet classification
using "Minkowski Functional"

Mihoko Nojiri & Sung Hak Lim

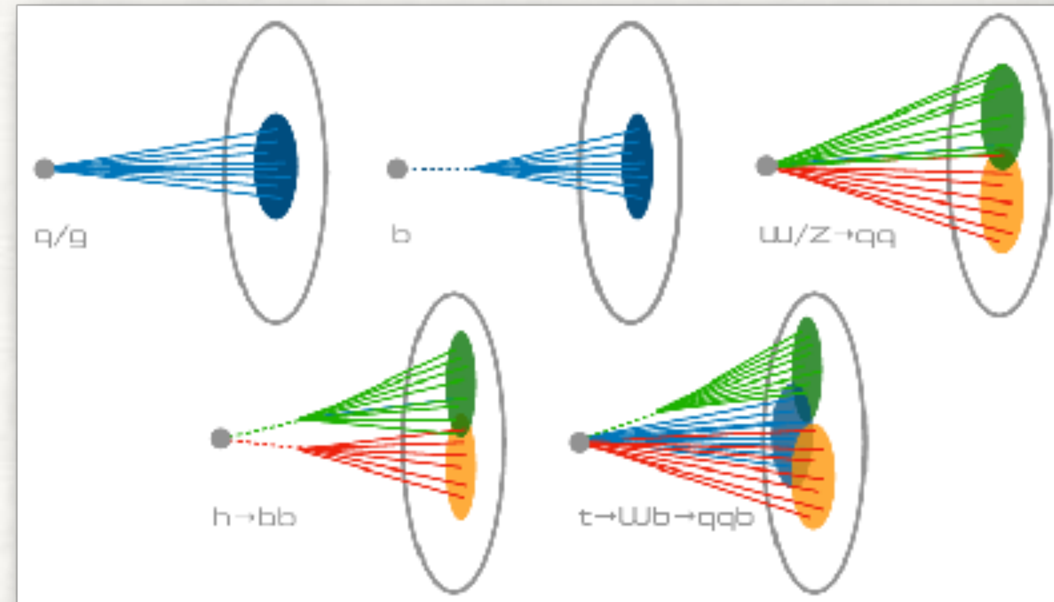
In preparation+ 2003.11787(JHEP), 1904.02092(JHEP)

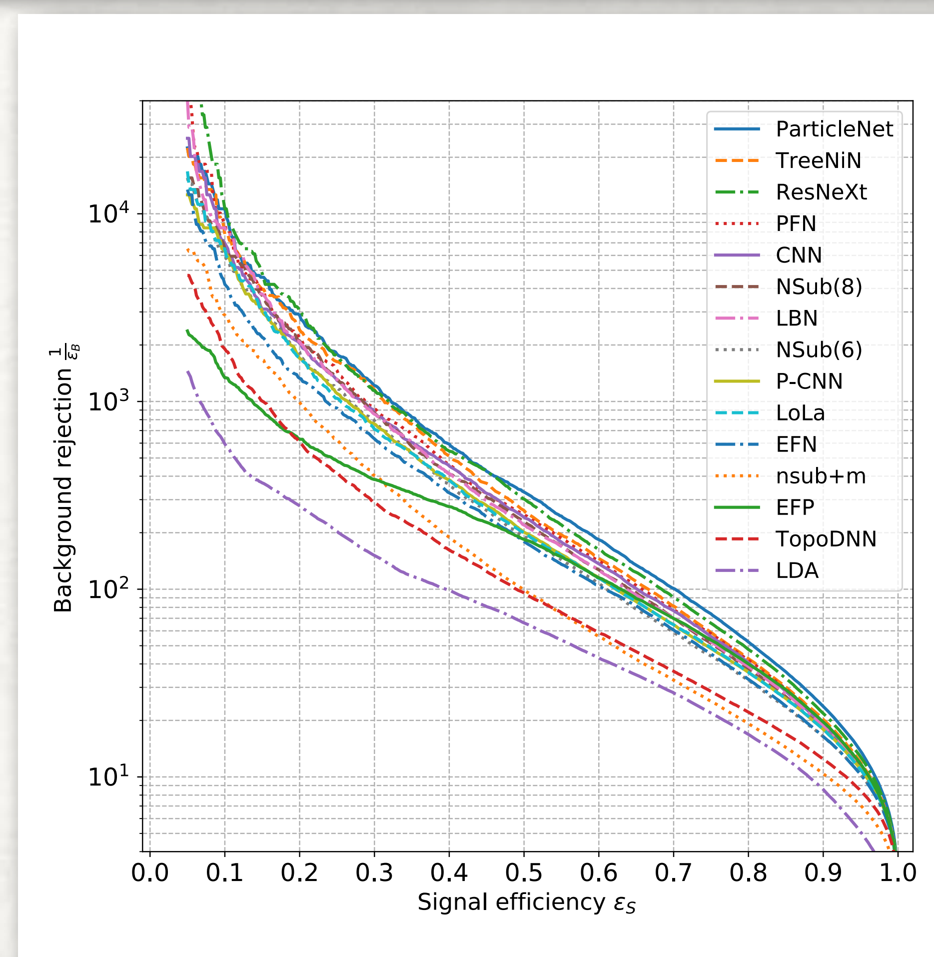
JET CLASSIFICATION AND DEEP LEARNING

JET CLASSIFICATION

- LHC -> HL-LHC -> FCC-hh (far future)
- Better sensitivity using ML (BDT)→ Deep Learning (Many architecture in market.)
  - QCD jet vs top, Higgs, W, new physics
  - Anomaly
  - High level inputs →low level inputs (Jet image) better

    performance. CNN, Recurrent neural network(RNN), particle net,,,

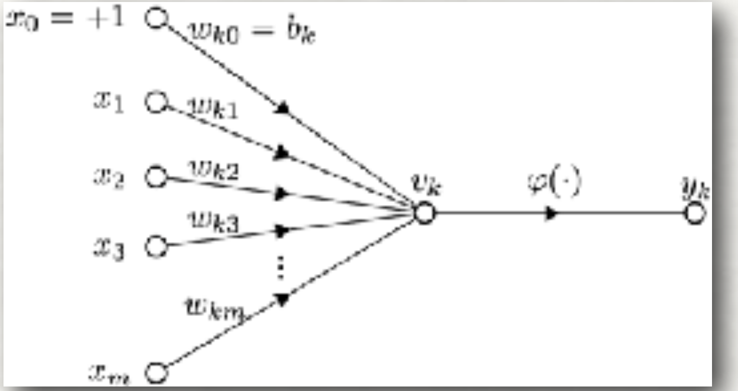


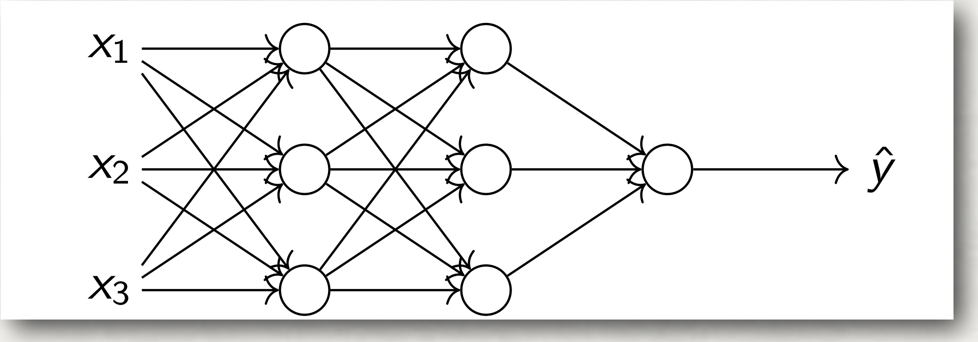


CLASSIFICATION USING DEEPLERNING

Structure

1. Inputs
2. Trainable parameters w, b
3. Activation function φ(wx+b)
4. Cross entropy minimization to get best classification





Signal (t=0) vs background (t=1)

$$L = -t \log y - (1-t) \log (1-y)$$

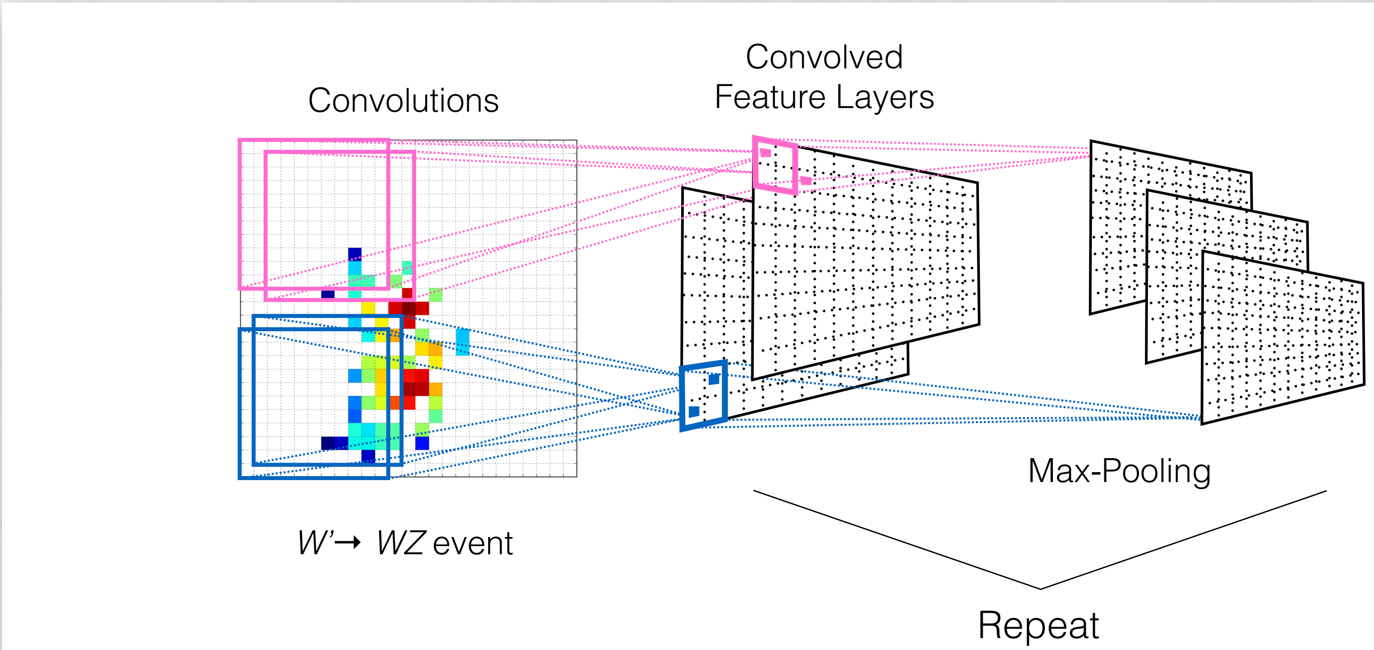
Merit Using NN:

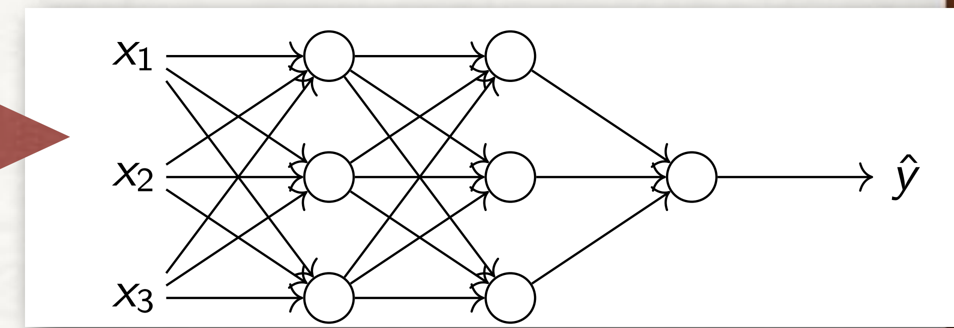
general function any reaction can be constructed

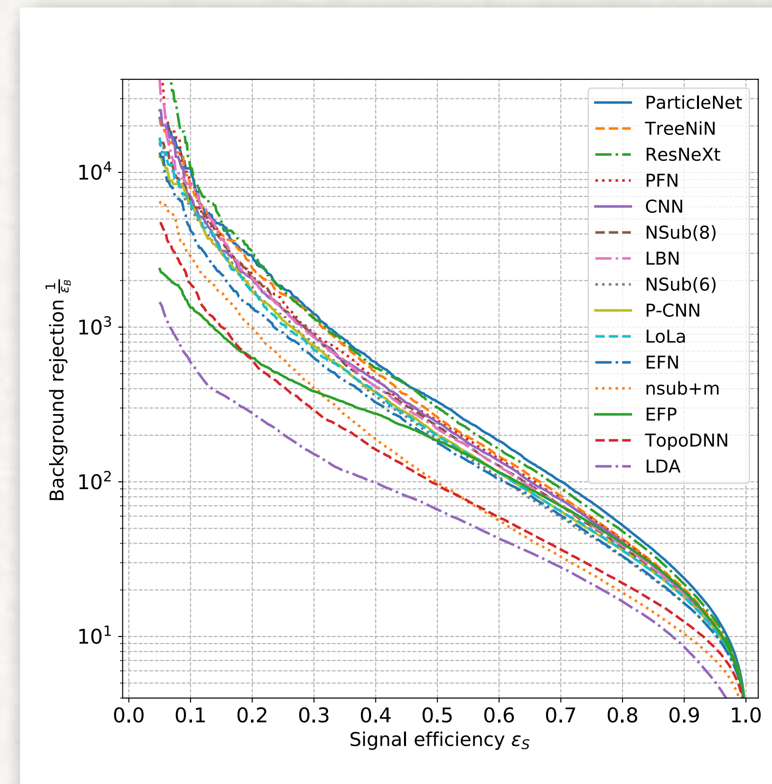
Fast with GPU!

But it is not easy to identify how they achieved good sensitivity.

CNN(CONVOLUTIONAL NEURAL NETWORK) AND TOP TAGGING

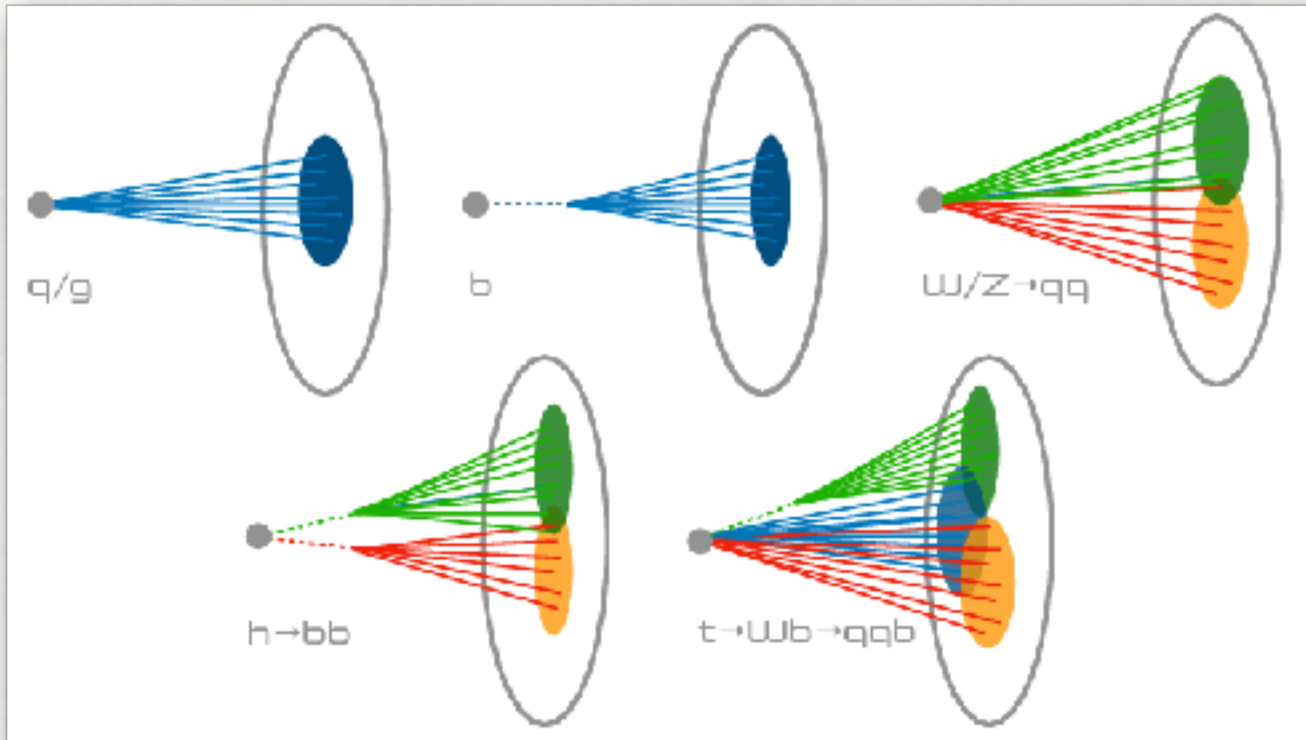


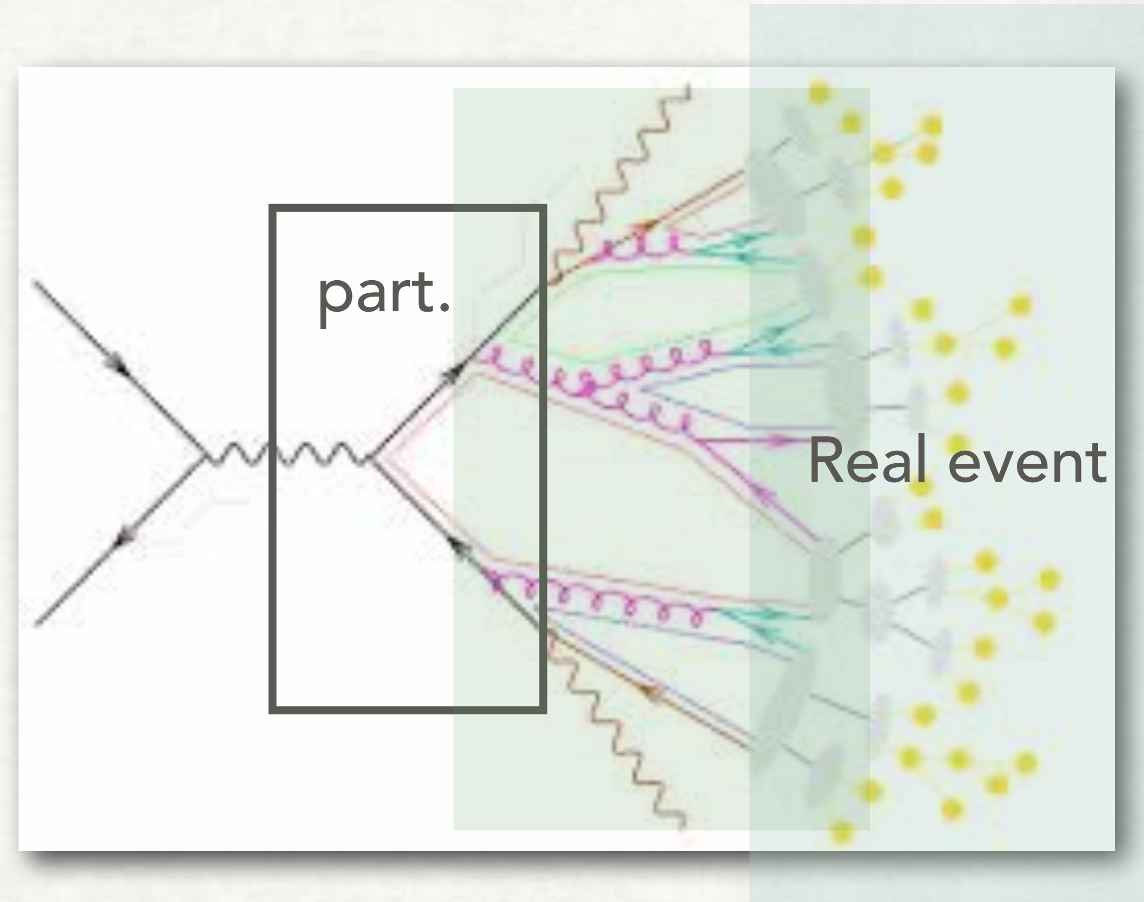




- Transfer image by NxN filter -> some cutoff (pooling) -> to find correlation.
- CNN, ResNeXT, Particle Net use correlation of particles and show simmer performance.
- **Why one is better than the other? What kind of event is excluded additionally. What is the key?**

TWO QUANTITY FOR JET CLASSIFICATION.





- IRC safe object: :subjet, Energy correlation(C-correlator)
- Objects Sensitive to Soft collinear splittings number of tracks, particles MC modeling Is bad (Pythia vs Hewig vs real data)
- Jet image contains both of them and jet classifier use it without prejudges
- Color coherence etc.. Soft particle distribution also have parent information

IT IS KNOWN THAT IRC SAFE OBJECT DOES NOT GIVE BEST

N-subjettiness MLP classification

$$\tau_N^{(\beta)} = \frac{1}{p_{TJ}} \sum_{i \in \mathrm{Jet}} p_{Ti} \min\left\{R_{1i}^{\beta}, R_{2i}^{\beta}, \ldots, R_{Ni}^{\beta}\right\} .$$

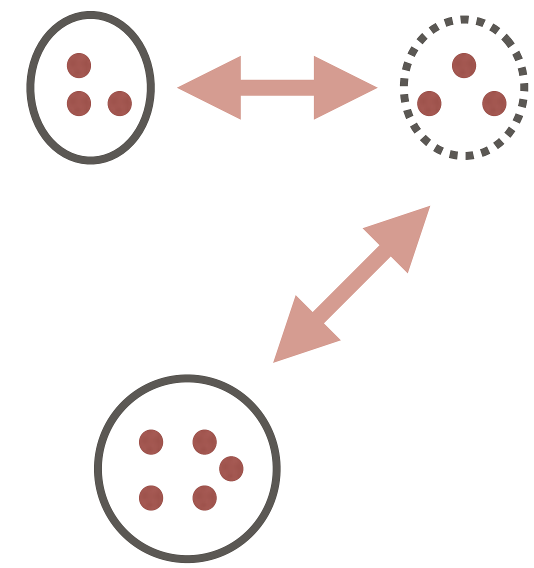
$$\left\{\tau_1^{(0.5)}, \tau_1^{(1)}, \tau_1^{(2)}, \tau_2^{(0.5)}, \tau_2^{(1)}, \tau_2^{(2)}, \tau_3^{(0.5)}, \tau_3^{(1)}, \tau_3^{(2)}, \tau_4^{(0.5)}, \tau_4^{(1)}, \tau_4^{(2)}, \tau_5^{(1)}, \tau_5^{(2)}\right\}$$

arXiv 1704.08249 Datta Larkoski

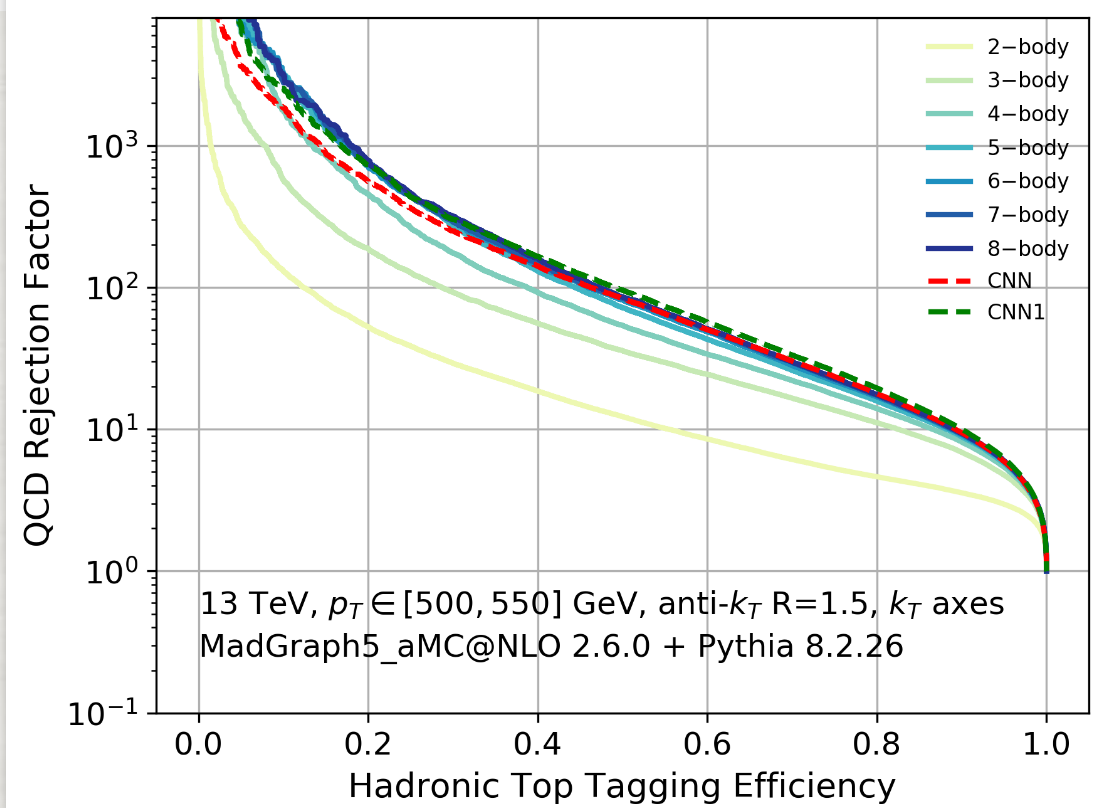
CNN vs N-subjettiness
Liam Moore et al 1807.04769

Need very higher order τ, does not looks interesting thing to do

τ2 LARGE T3 SMALL

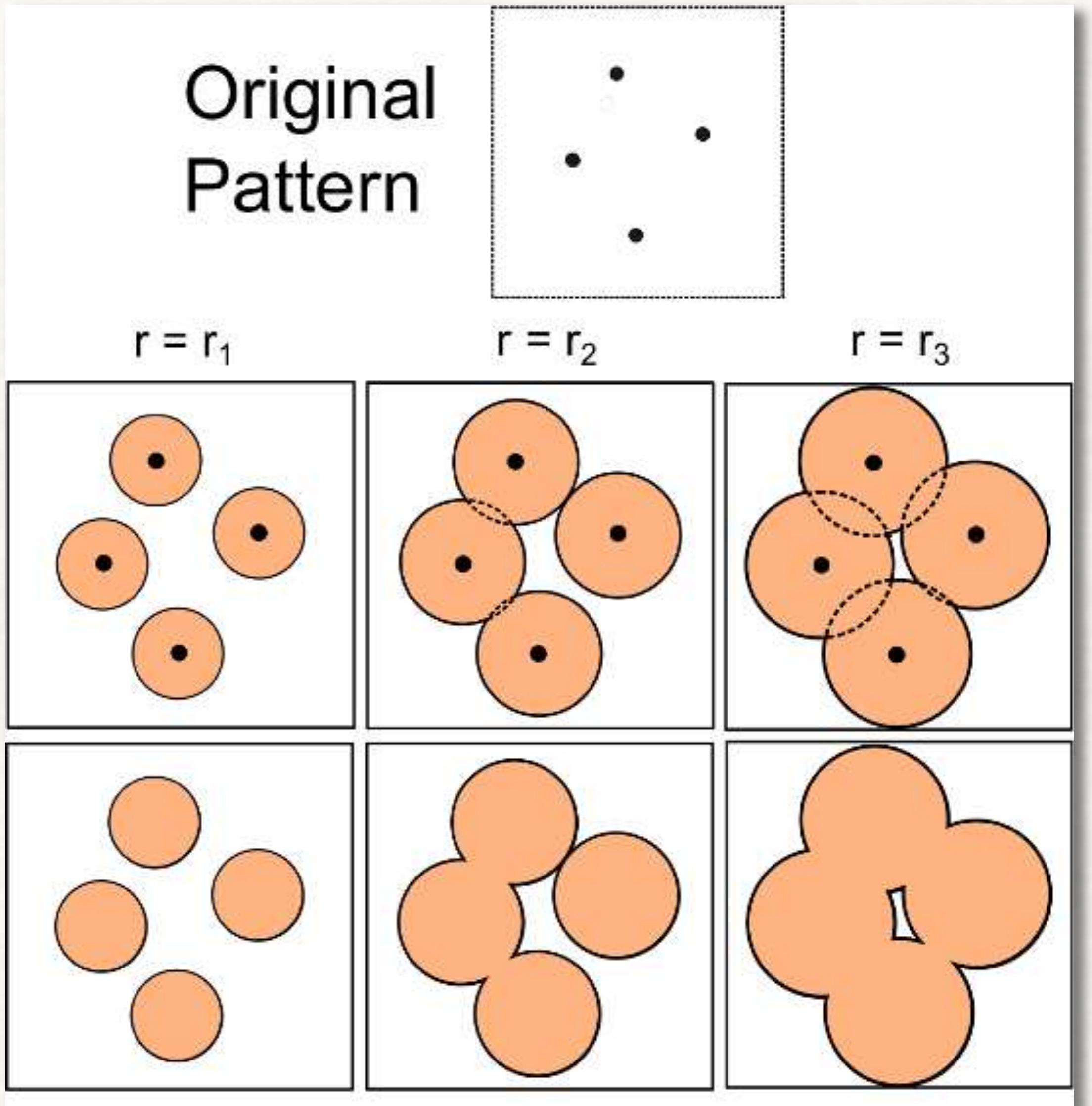


Coaster 2



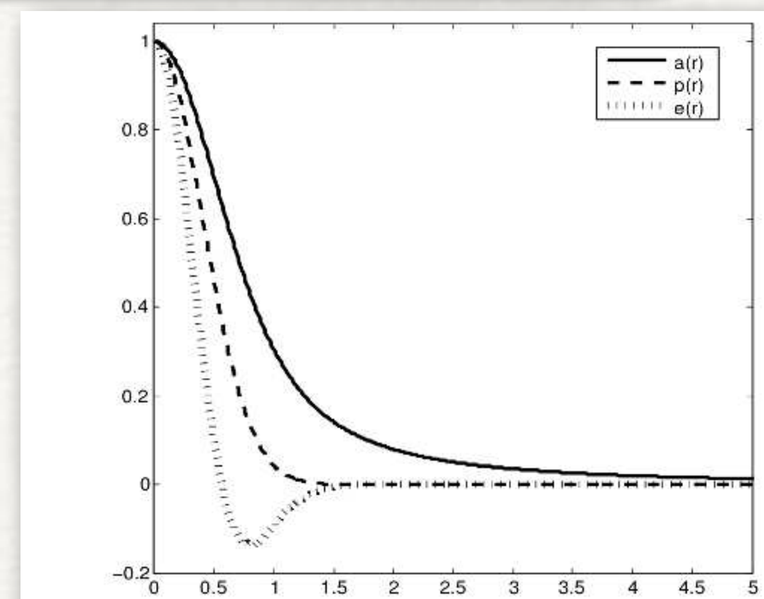
MINKOWSKI FUNCTIONAL AND JET CLASSIFICATION

"MINKOWSKI FUNCTIONAL MAP" POINT DISTRIBUTION TO REAL FUNCTIONWITHOUT LOSING INFORMATION



(a)

Area	$a(r) = \dfrac{A(r)}{\pi N r^2}$
Perimeter	$p(r) = \dfrac{P(r)}{2\pi N r}$
Euler characteristic	$e(r) = \dfrac{\chi(r)}{N}$



(b)

In N dim → only N+1 functional that satisfy

$M(B_1 \bigcup B_2) = M(B_1) + M(B_2) - M(B_1 \bigcap B_2)$ and rot and trans invariant.

APPLICATION OF MF

Statistical Physics

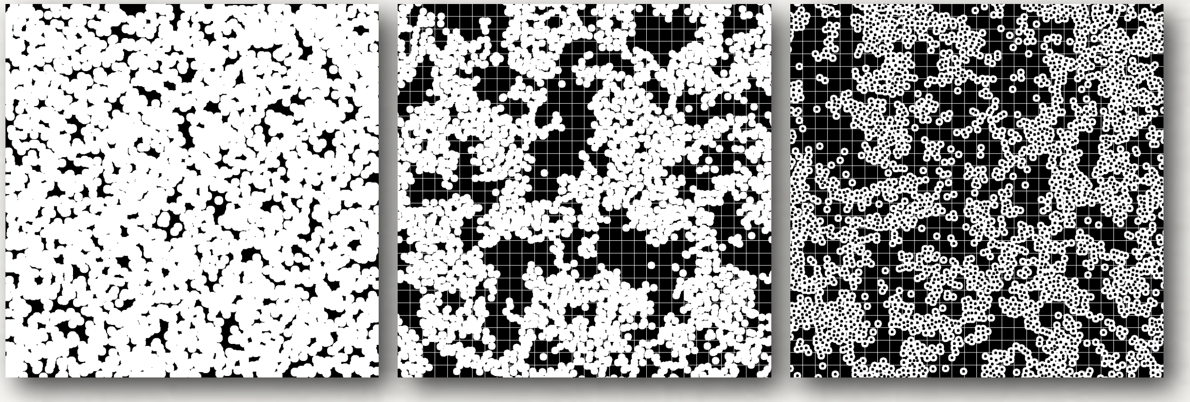
Left Porous medium

Middle: microemulsion

Right : Colloid

Occupation V, Surface(S) → material physics

Mecke and Stoyan (2000)



- Astrophysics : star and galaxy distribution, simulation study, non-Gaussinaity of CMB, weak lensing..

Powerful to quantitatively describe point distribution

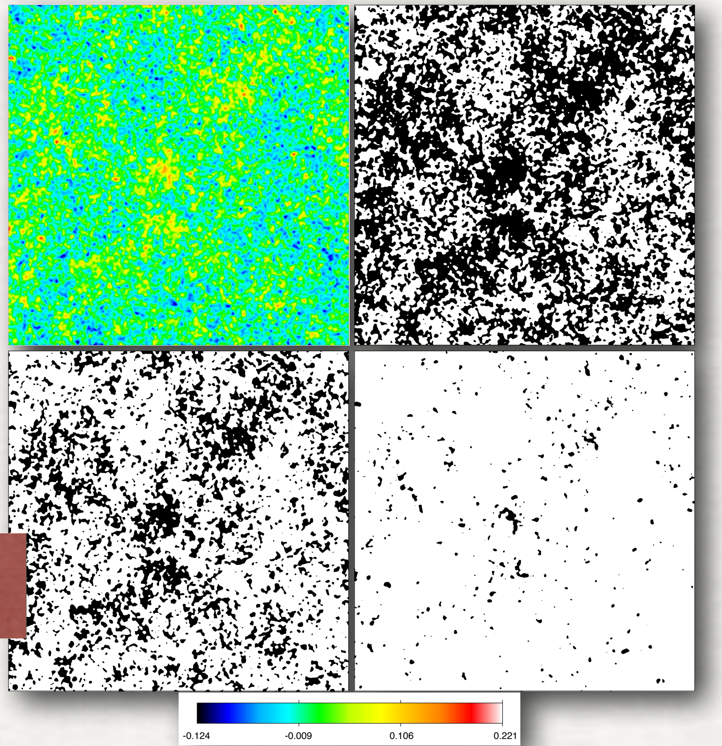
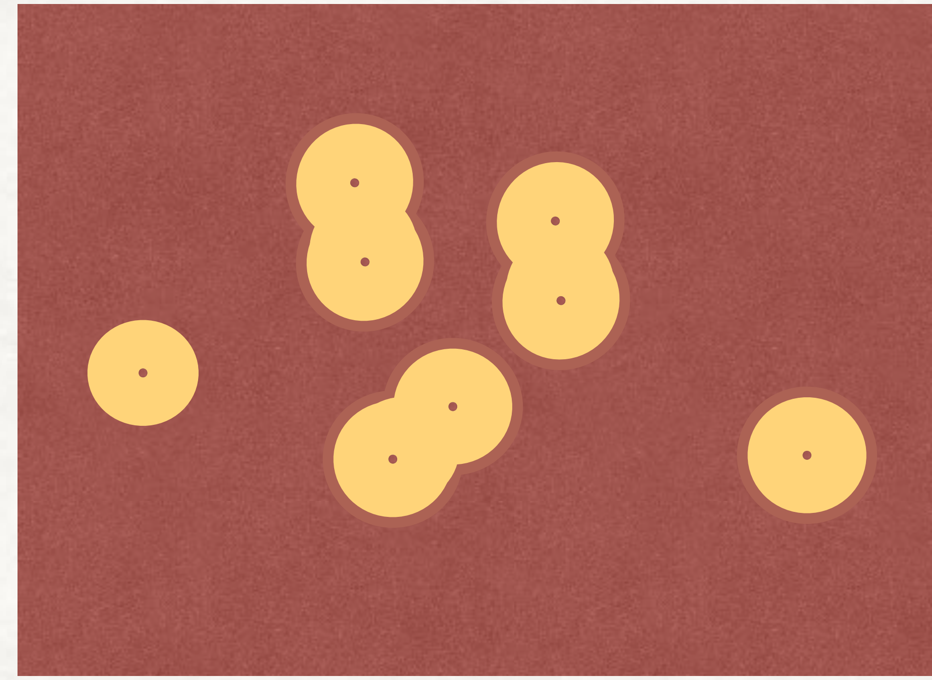


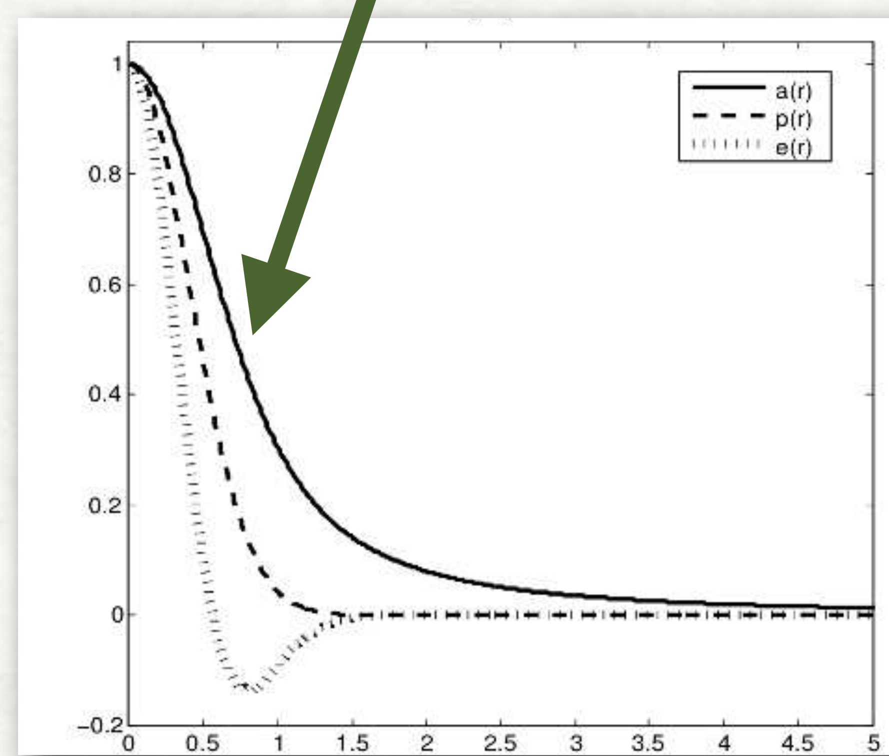
FIG. 1: *Top left panel: example of a simulated 12-square-degree convergence map in the fiducial cosmology, with intrinsic ellipticity noise from source galaxies and $\theta_G = 1$ arcmin Gaussian smoothing. A source galaxy density of $n_{gal} = 15/\text{arcmin}^2$ at redshift $z_s = 2$ was assumed. Other three panels: the excursion sets above three different convergence thresholds κ, i.e. all pixels with values above (below) the threshold are black (white). The threshold values are $\kappa = 0.0$ (top right), $\kappa = 0.02$ (bottom left), and $\kappa = 0.07$ (bottom right). The Minkowski Functionals V_0, V_1, and V_2 measure the area, boundary length, and Euler characteristic (or genus), respectively, of the black regions as a function of threshold.*

Kratochvil 1109.6334 Proving Cosmology with Weak Lensing Minkowski Functional s

APPLICATION TO JET PHYSICS

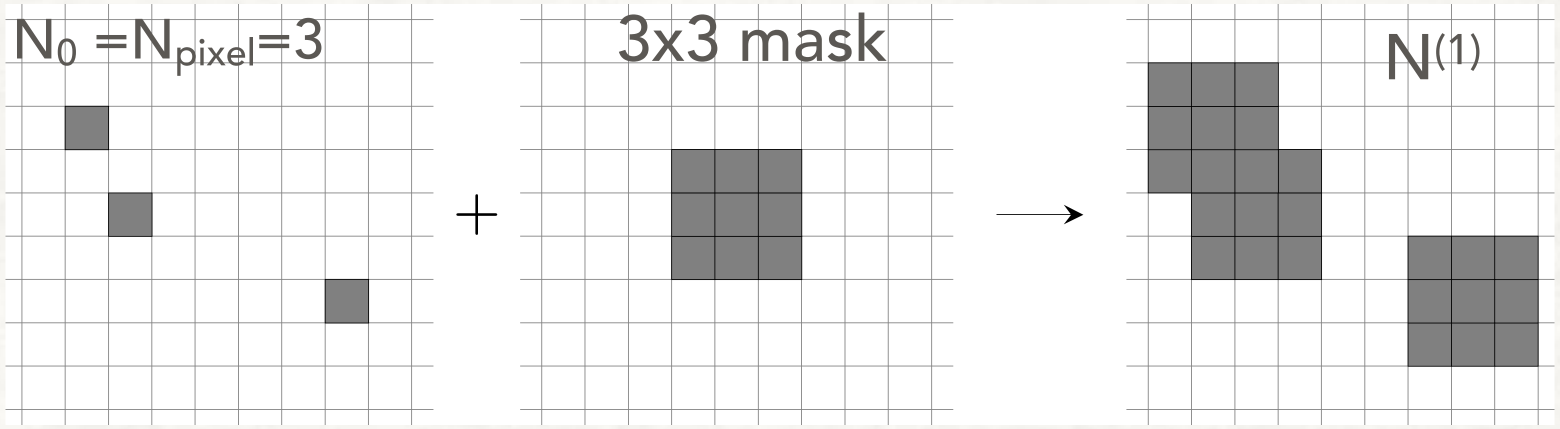
- only good things.
  - MF treats distance between points equally.
  - No loss of information
  - Pixel by pixel fluctuation of CNN input reduces to 1/Npixel fluctuation of MF. This improve loss function minimization.
  - The concept common to modern jet algorithm. (Jet area)
  - MF is stable in collinear splitting because it does not double count overlapping area.





IMPLEMENTATION

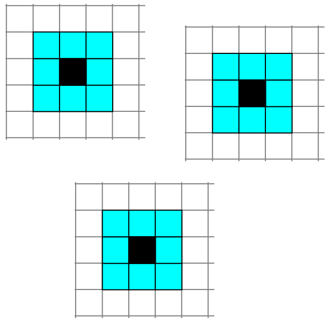
- Jet image prepare mask of size 3x3 , 5x5, to all active pixel



- What is the information in, for example, N_1/N_0

Isolated points

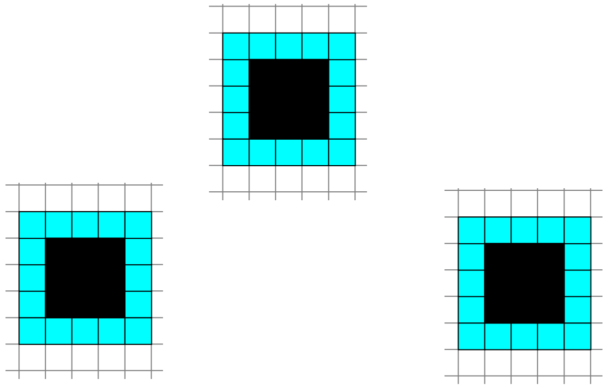
$N_1/N_0 = 9$



VS

all pixel appear in 3x3

$N_1/N_0 = 16/9 = 1.78$

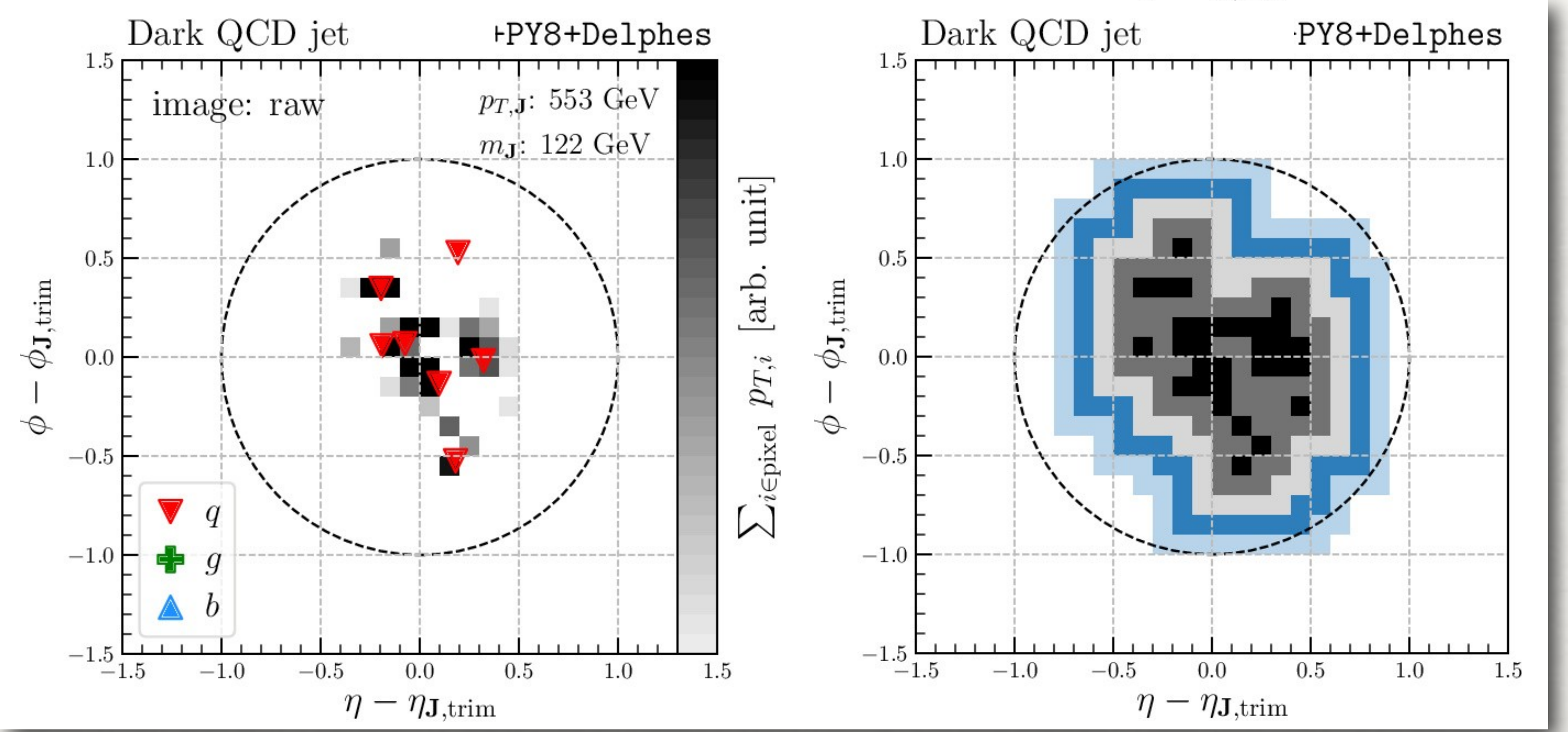


A CASE OF DARK JET

Lim, Nojiri   in preparation

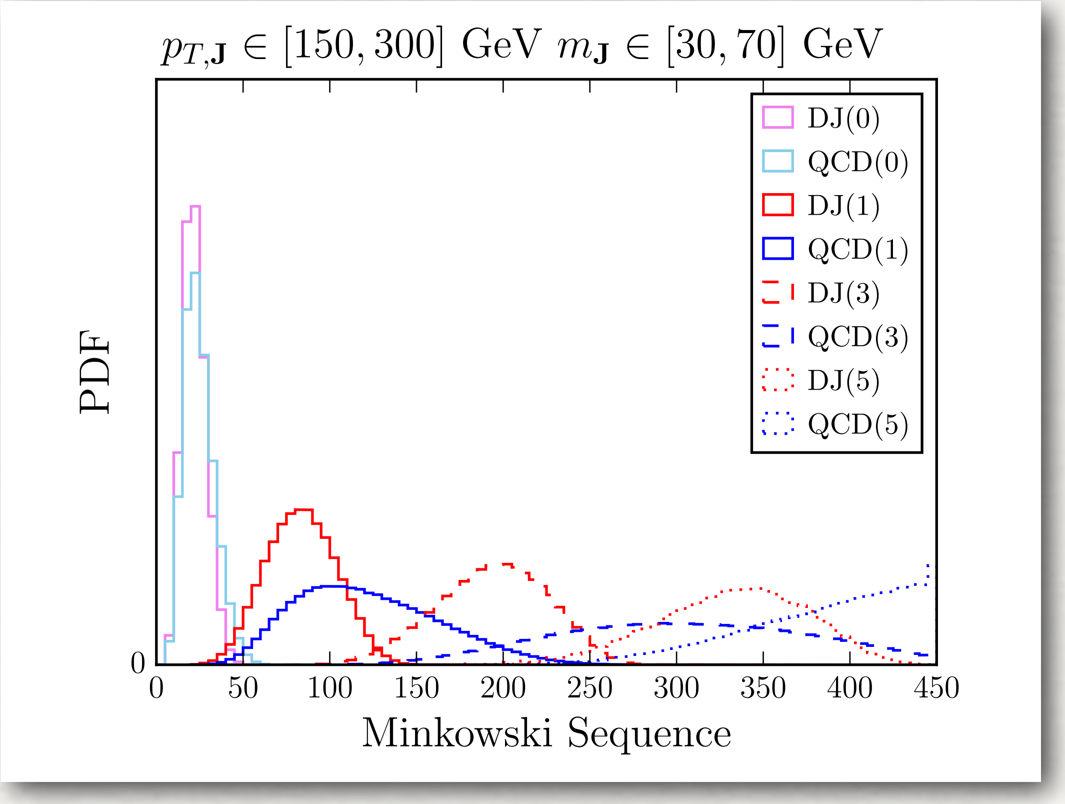
- Dark Jet   pp →Z′ → q_D $q_D$bar→dark Parton shower → ρ_{diag}→qqbar
- Multiple “color singlet” cluster in the jet

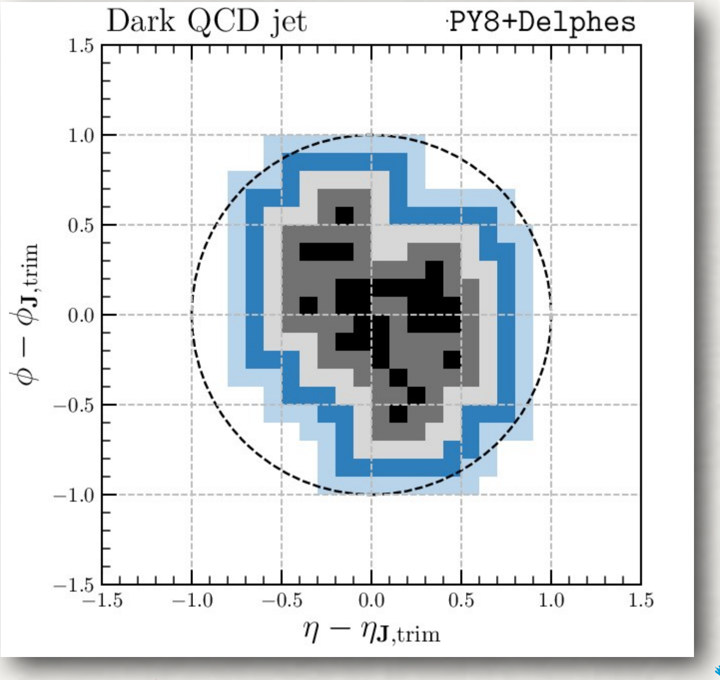
m_ρ=20GeV



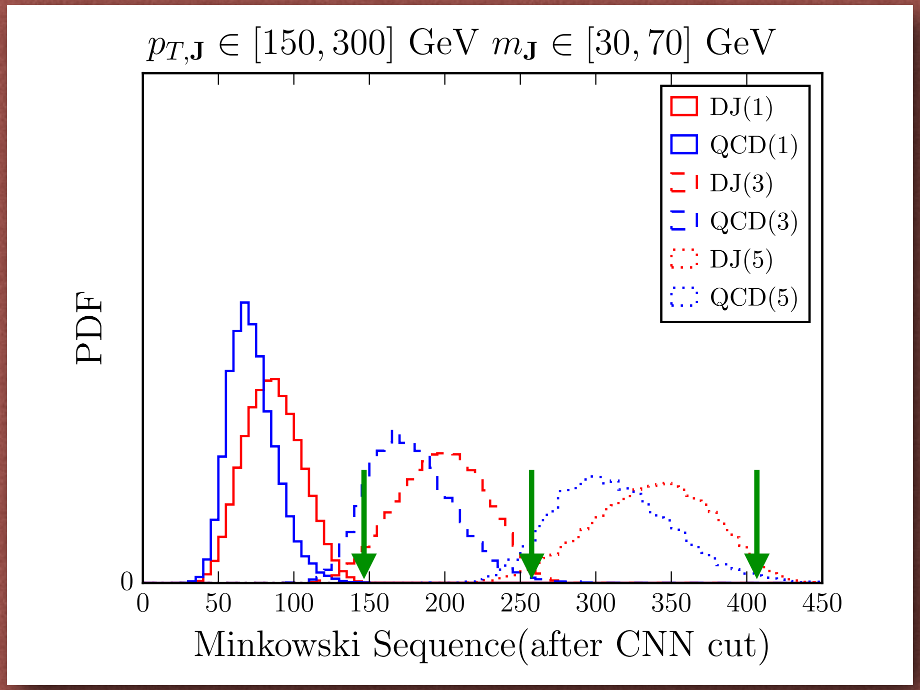
WHAT CNN TRAINING DOES?

Original distribution





Cut using CNN



Smaller MF endpoint suggests Compact soft activities

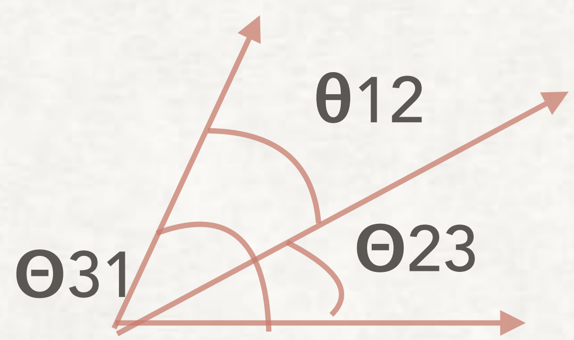
CNN allowing 10% rejection of signal

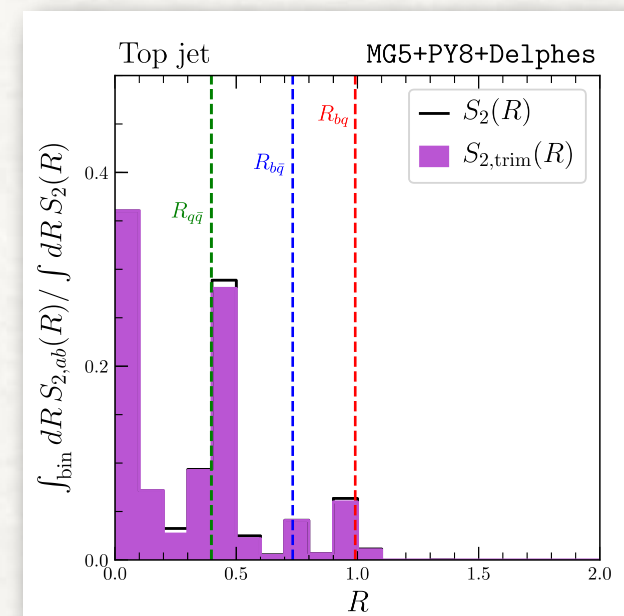
IMPROVING JET CLASSIFICATION WITH MF

- Training MLP with MF +global variable(Jet mass, jet pT, trimmed jet PT, the leading subjet and their mass) already reject significant events for Dark jet

- In addition we introduce **relation network(RN)**

- Namely, **aggregated two point energy correlation** of jet and trimmed jet S_2 (θ) =Ei Ej δ(θ−θij)





  - S2 is C- correlator and IRC safe **(Tkachov hep-ph 9601308)** and $EFP_2^{(n)}$ can be reproduced from S2. Lim, Nojiri 1807.03312, Chakraborty, Lim Nojiri 1904.02092

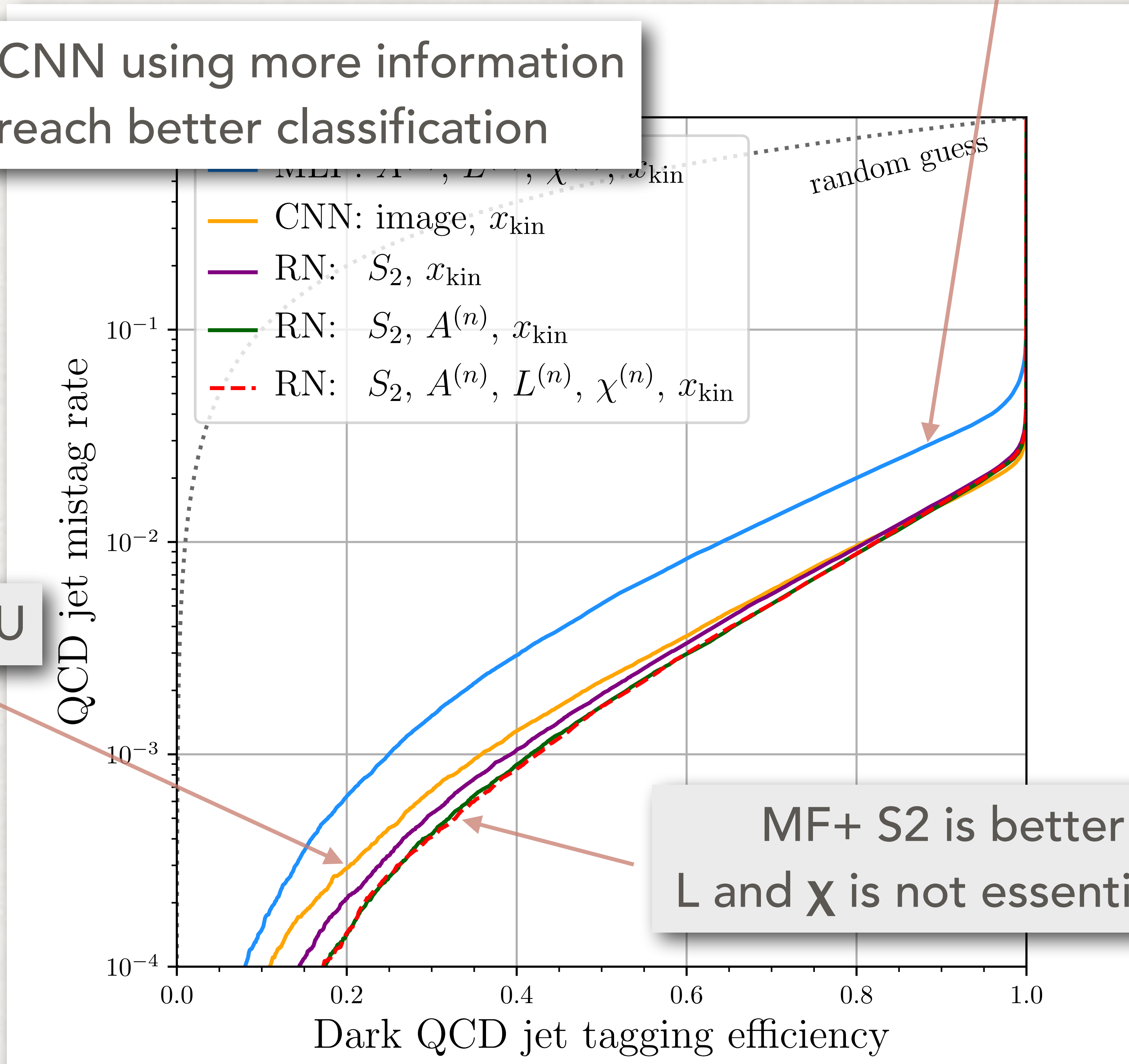
- S2[Leading subjet] x [its counter J/J1]   and S2[J/J1]x [J/J1]: For top, three point correlation is important

RESULTS FOR DARK JET VS QCD

Jet PT 150~300GeV Jet mass 30~ 70GeV

MF only reject lots of QCD jet

Even though CNN using more information
RN+ MF reach better classification



CNN using ReLU

MF+ S2 is better than CNN
L and χ is not essential for this case

RN + MF also improve Top vs QCD (but improvement is smaller)

GUESSING WHY CNN IS WORSE

- Convolutional neural net(CNN) can access MF if the distributions of QCD and DJ are different.

Realization of MF by 2x2 filters

v: lookup table from 2x2 partition to 3 vectors that sum up to MF

$$v\,(f_i) \rightarrow (A_i,\, P_i,\, \chi_i)\ 、\ (A, P, \chi) = \Sigma_i\, v(f_i)$$

However

1. Jet image is weighted by energy.

2. To access MF, you need to construct the function with no energy dependence (function of binary image)

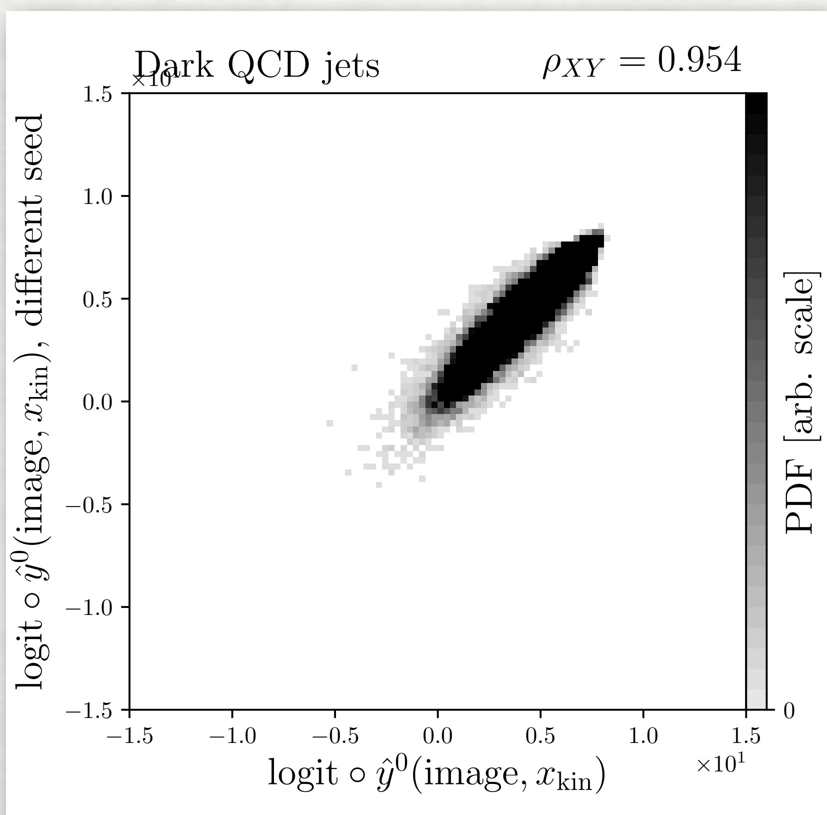
Table 1. Look-up table for Minkowski functionals.

Conf.		A	P	χ	Conf.		A	P	χ
1		0	0	0	9		1/4	1	1/4
2		1/4	1	1/4	10		1/2	2	−1/2
3		1/4	1	1/4	11		1/2	1	0
4		1/2	1	0	12		3/4	1	−1/4
5		1/4	1	1/4	13		1/2	1	0
6		1/2	1	0	14		3/4	1	−1/4
7		1/2	2	−1/2	15		3/4	1	−1/4
8		3/4	1	−1/4	16		1	0	0

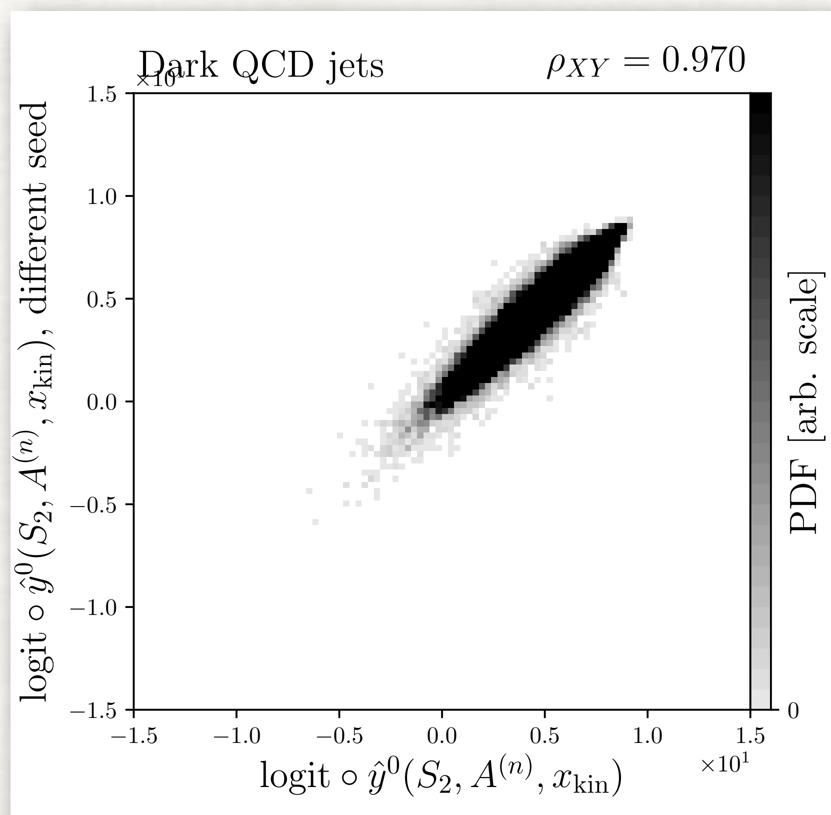
SCORE DISTRIBUTION

- Loss function minimization suffers from fake minimum.
- Systematics estimate by changing event ordering in batch training: Correlation among the same model is typically above 0.9. CNN vs RN+MF is 0.79.

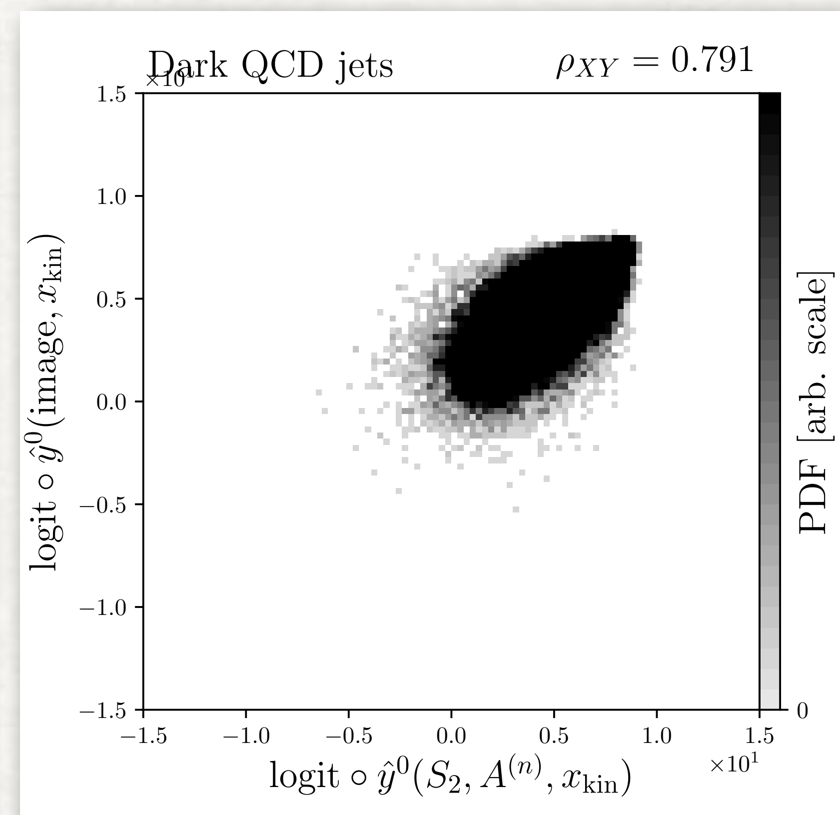
CNN vs CNN



RN vs RN

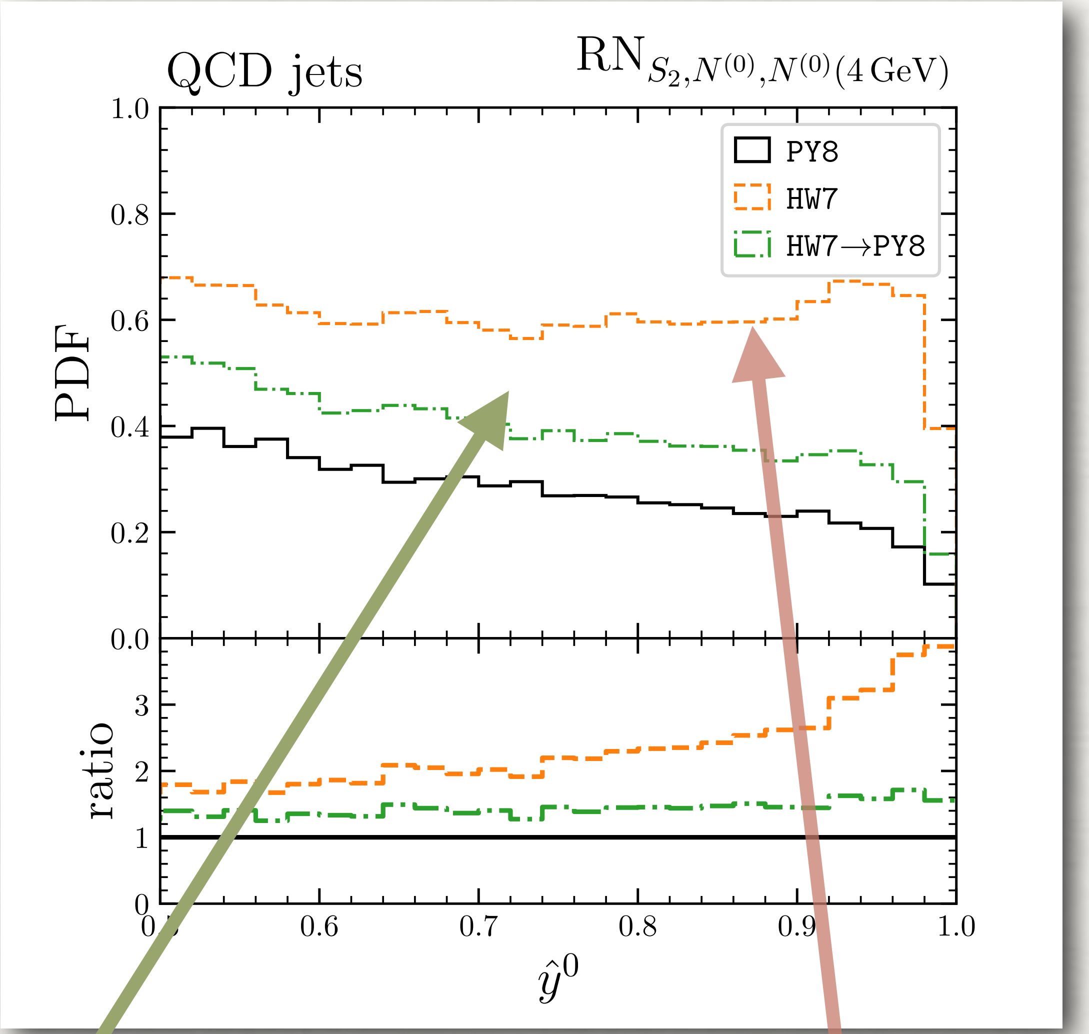


CNN vs RN



MF FOR MC CALIBRATION

- Herwig, Pythia, Sherpa: MC does not agree on soft particle distributions, especially for QCD jets.
- Improving the model using data takes time.
- Event reweighing via MF rather than bin by bin event distribution: IRC safe distribution is not affected and improve the overall agreement in "classification"



Pythia trained classifier classifying Herwig QCD jet

Reweighting to reproduce A(0)

SHORT SUMMARY

- Minkowski functional is good tool to describe the n dim distributions of featureless points, and it works in Jet Physics too.
- Dark Jet vs QCD: CNN discover MF (without being told)
- MF + RN is better than CNN although(and because) it uses only part of the jet image. Reducing fluctuation by aggregation is the key.
- Application to other physics? (Such as displaced vertex, other detector such as water tank, cosmic shower) …

BACK UP : OUR RELATION NETWORK