

DARK JET の深層学習による分類と
ジェットの幾何学
(ミンコフキー汎関数を中心にして)

KEK 野尻美保子 & Sung Hak Lim (KEK→ラトガス)

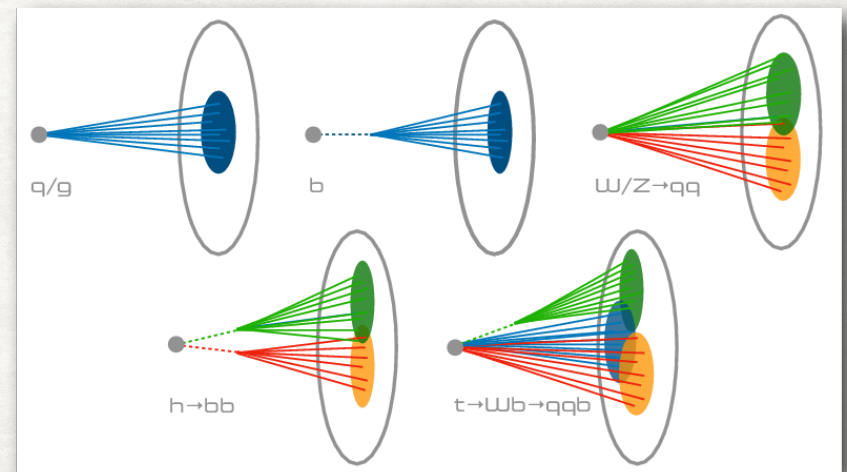
このトークのまとめ

- ジェットの深層学習の問題点(摂動的な量と非摂動的な量を区別せずに使っている)
- ミンコフスキー汎関数(MF)を使って、非摂動的なエネルギーの低い粒子の分布に特化した測定量を構築
- 畳み込みニューラルネット(CNN)が
 - MFを自律的に発見している例 (Dark Jet)
 - MF が見つけられずに中途半端になっている例(top ジェット)
 - MF を実験データから補正するには

ジェットの深層学習と問題点

ジェットの深層学習

- 今後のLHC の物理 高輝度化 (HL-LHC) 高エネルギー化 (ずっと先)
- 感度の向上への期待
 - 高エネルギーのtop, Higgs, W からくる粒子の選別 (boosted object) バックグラウンドはQCDジェット
 - アノーマリの選別
- 機械学習(BDT) から深層学習
- 様々なアルゴリズムの提案
 - 高次の量の利用→低次の量(Jet image など)
MC レベルではよりよい結果を示す。例えば top Higgs vs QCDジェットなど



分類問題と深層学習

深層学習の構成

1. 特徴を自動的に抽出するための、調整可能なパラメータスペース
2. 活性化関数: 決定のための非線型応答
3. 損失関数 (交差エントロピー)

シグナル (t=0) vs バックグラウンド (t=1)

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

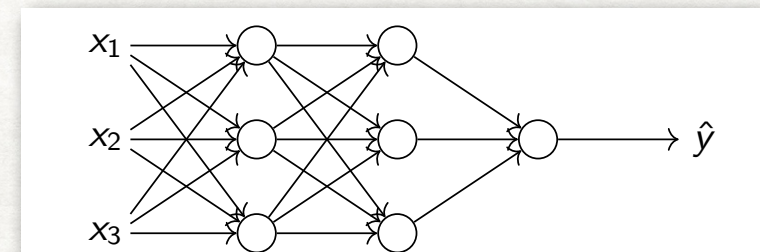
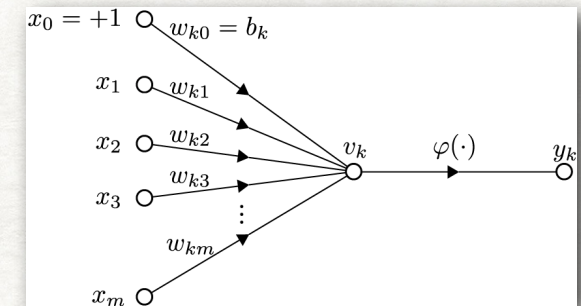
深層学習が得意なこと

任意の応答を記述可能

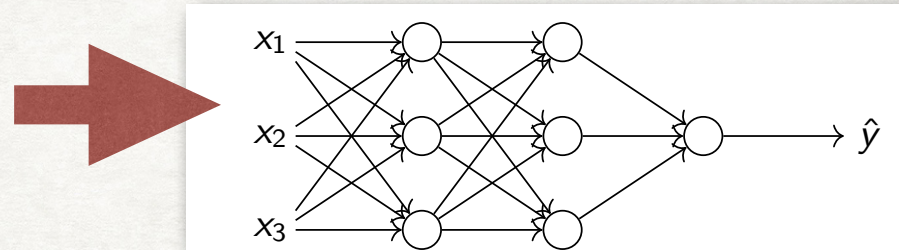
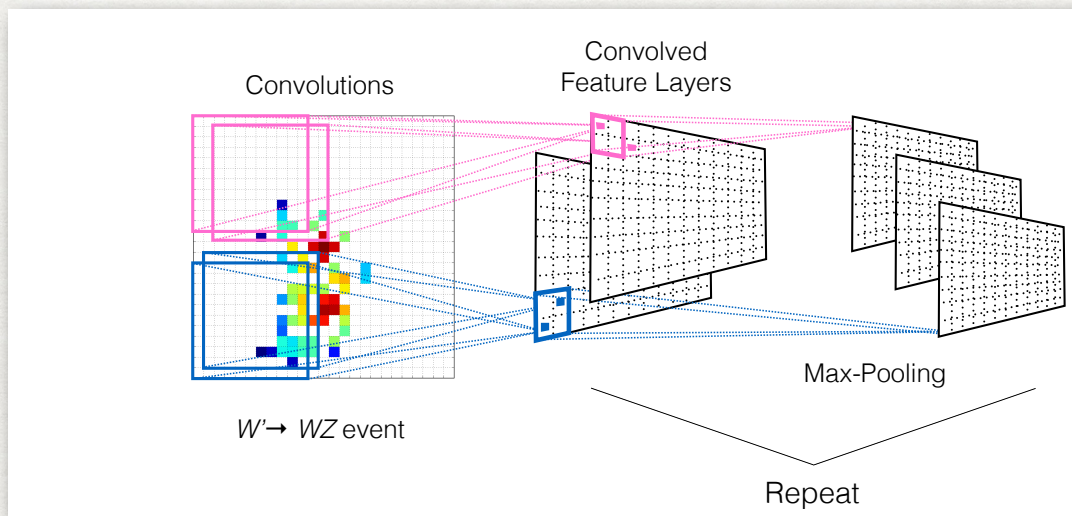
高速化が可能

イベントを少しずつまとめて学習させ、 w, b を損失関数が最小になるように調整する。

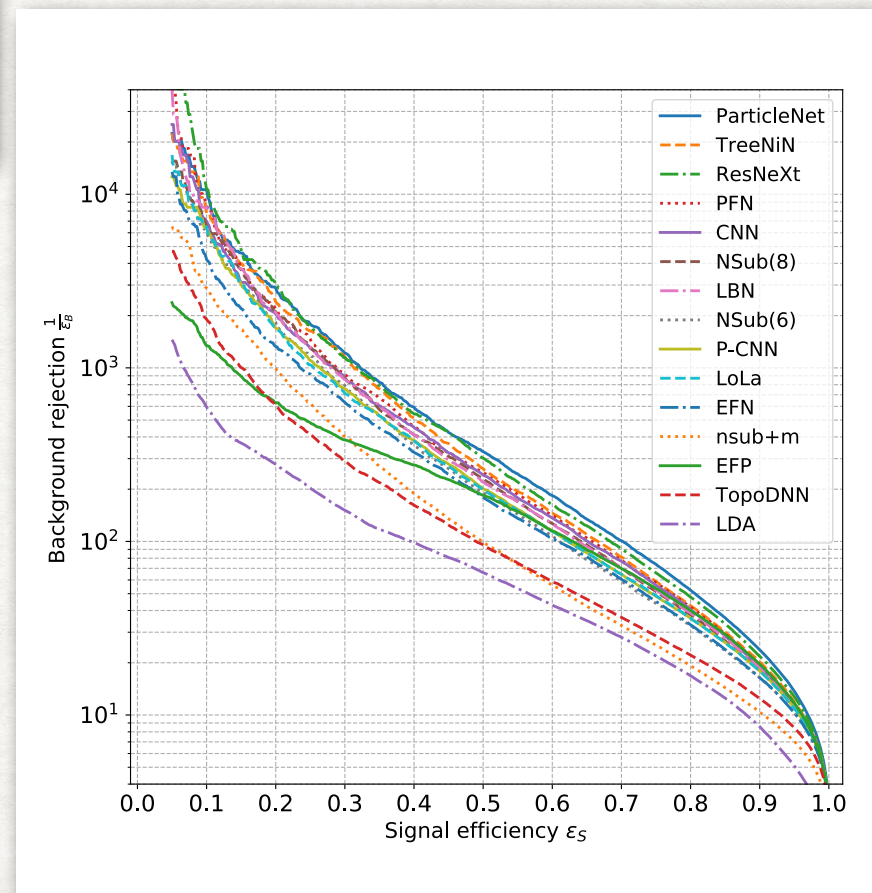
様々な分類を効率的にこなすが、特徴量を抽出することが難しい



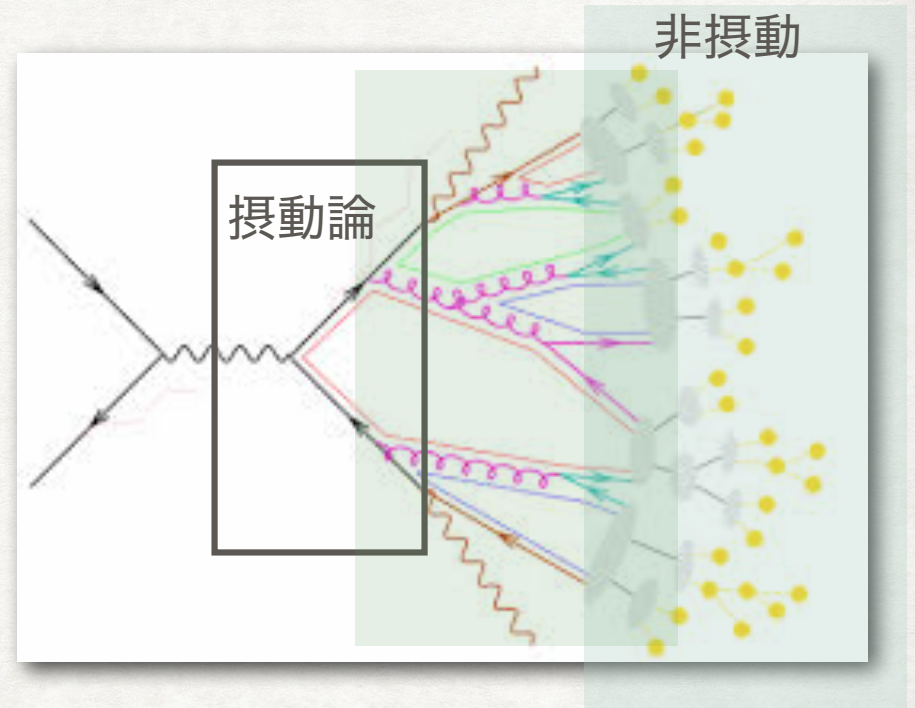
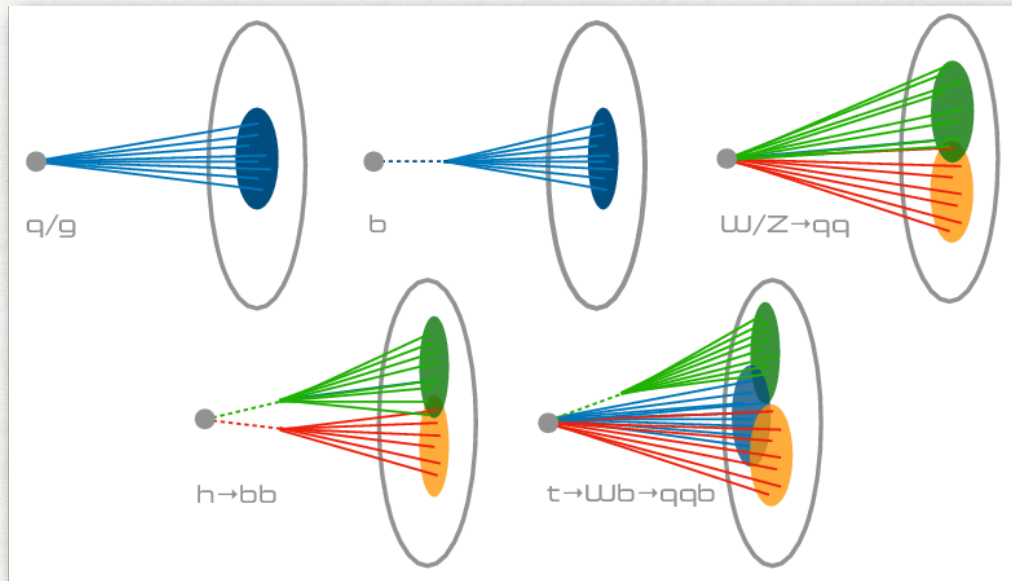
CNN(畳み込みニューラルネットワーク) とTOP TAGGING



- 小さい $N \times N$ のフィルタから特徴量の抽出を繰り返す。
- 多くの実装: CNN, ResNeXT, Particle Net... は大体同じパフォーマンスを示す。
- イベントごとの比較が不在、なぜ良いかも不明



深層学習の問題点



- ジェットの中に2つの量がある。
 - IRC safe object: soft or collinear emission に対して安定な量: subjet
 - Soft collinear に対して不安定な量 number of tracks, particles MC modeling に大きな差 (Pythia vs Hewig vs 実験データなど)
- ジェットイメージには全てが優先順位なくに入っている。

ハードな量だけの深層学習はベストな結果を出さない

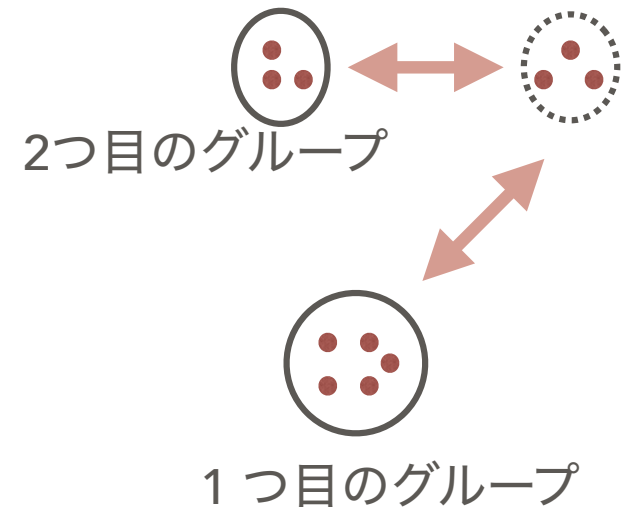
N-subjettiness を使ったNN

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

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

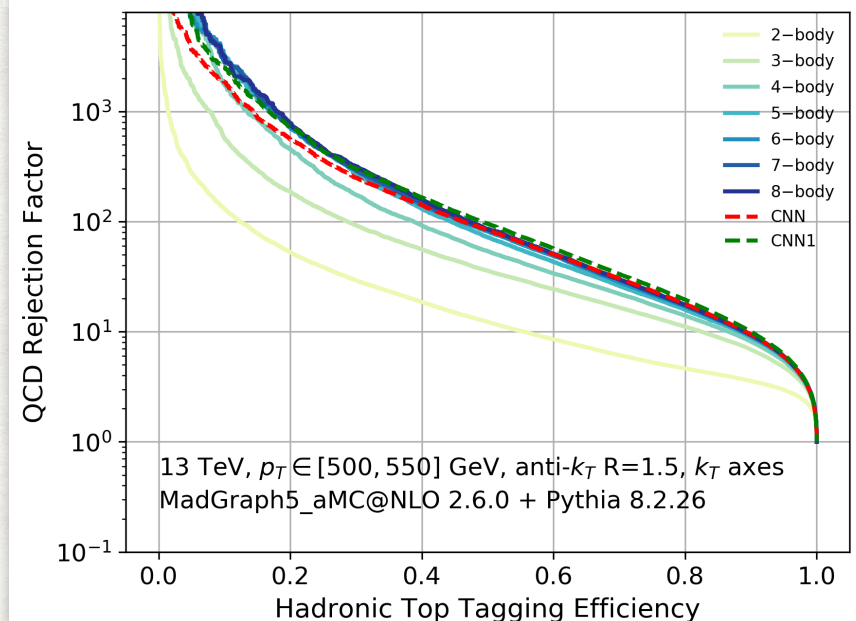
arXiv 1704.08249 Datta Larkoski

T₂テンション→T₃ で解消の例



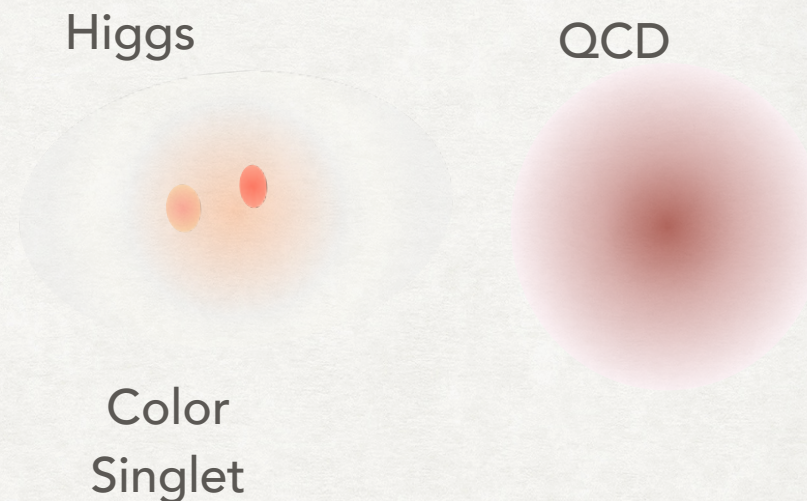
CNN との比較 Liam Moore et al 1807.04769

DNN に高次の N-subjettiness 入力してもCNN の効率に到達しない。残念



SOFT な物理 カラーコヒーレンス

- カラーコヒーレンス: Higgs のようなカラーシングレットな粒子の作るジェットは、大きい角度にハドロン粒子を放出しないはず。
 - W, H, Z 粒子束が独立する傾向
 - QCD より遠方まで粒子を飛ばす傾向
 - クォーク (どちらかと言うと) 芯がある vs グルーオン より広い N_q/N_g は理論的に計算できる量
- Event generator : Long Distance の効果を適当なモデルを作って、記述する。モデルのパラメータは実験データで補正する。

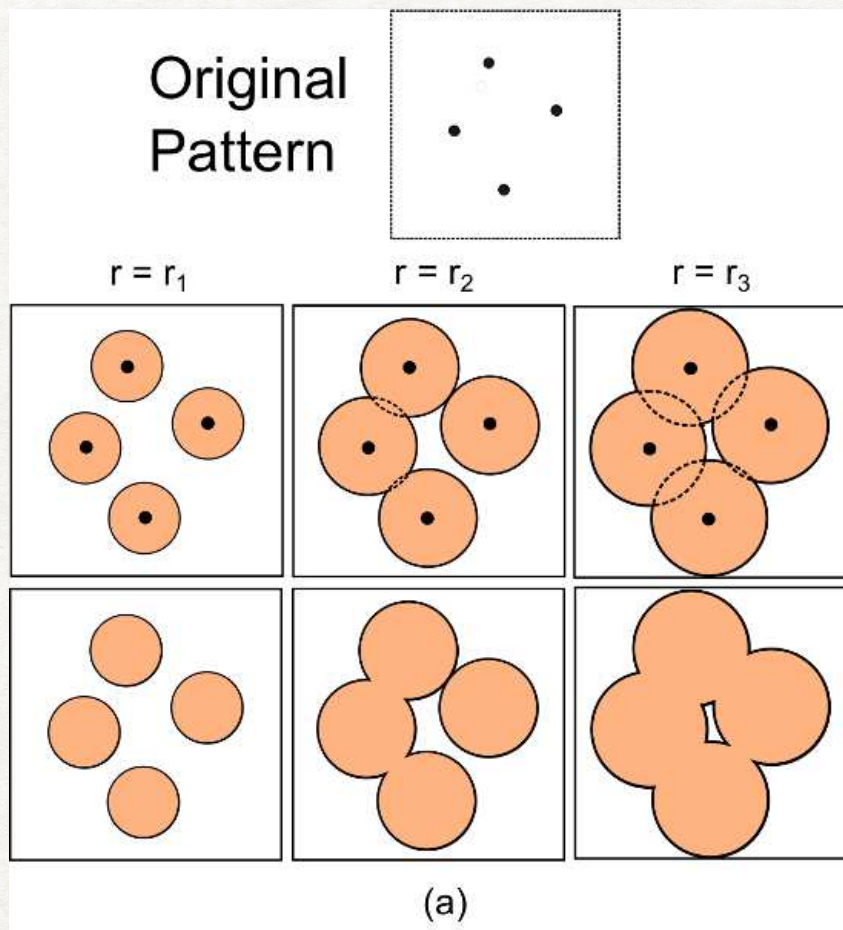


Trimming, soft drop, Iterated soft drop

深層学習で、エネルギーが低いハドロンの分布はどう影響するか

ジェットのミンコフスキー汎関数に
よる粒子分布の定量化

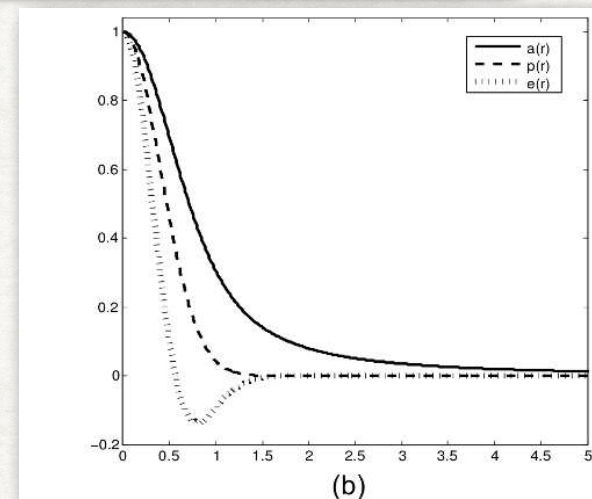
ソフトな粒子の分布を定量化するミンコフキー汎関数



面積 $a(r) = \frac{A(r)}{\pi N r^2}$

境界の長さ $p(r) = \frac{P(r)}{2\pi N r}$

オイラー標数 $e(r) = \frac{\chi(r)}{N}$



2次元上では 面積、周囲の長さ、オイラー標数(r 依存)で
点分布が記述される

MF の物理応用

統計物理

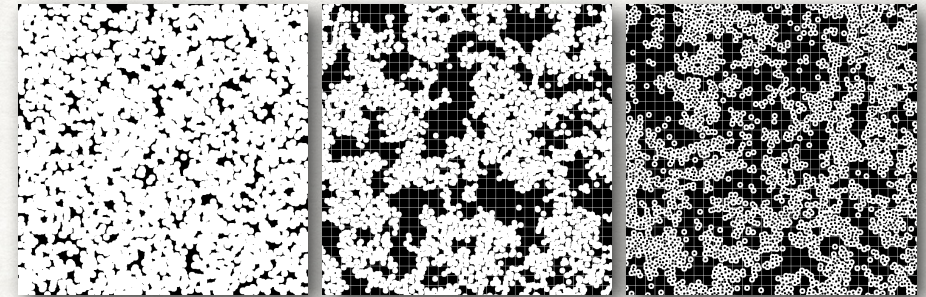
左 多孔質体

真ん中: 微乳濁液

左 コロイド

体積の占有状況 V , 表面の大きさ S 等に依存して物性が変わる

図は Mecke and Stoyan (2000)



- 天文: 星の分布の定量化、銀河分布、シミュレーション結果の定量化、non-Gaussianity of CMB, weak lensing..

点の集まりの意味を定量的に表現する時に使う

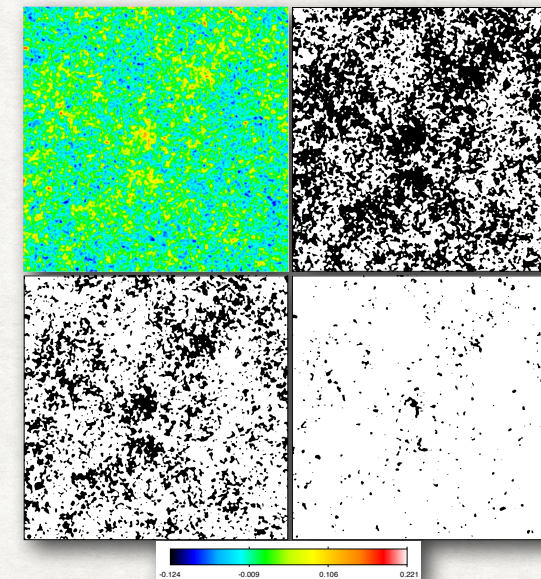
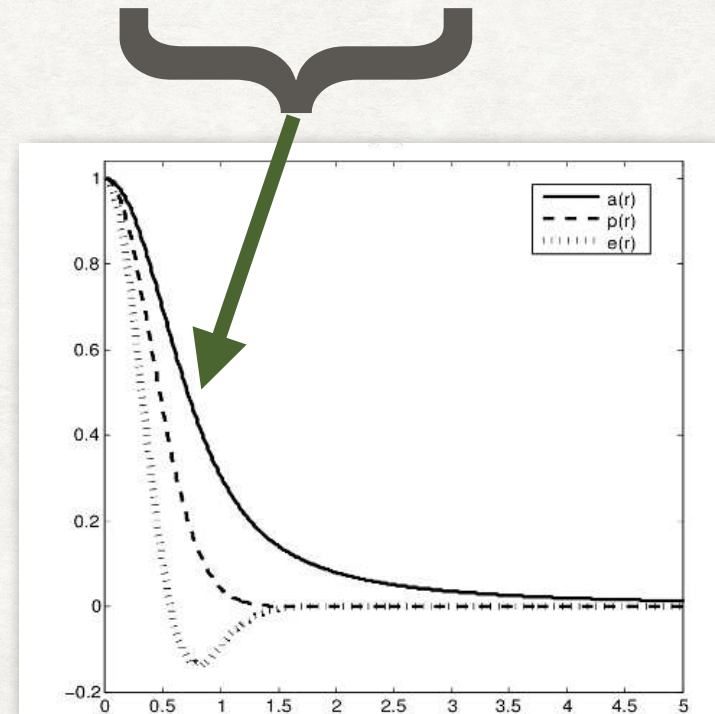
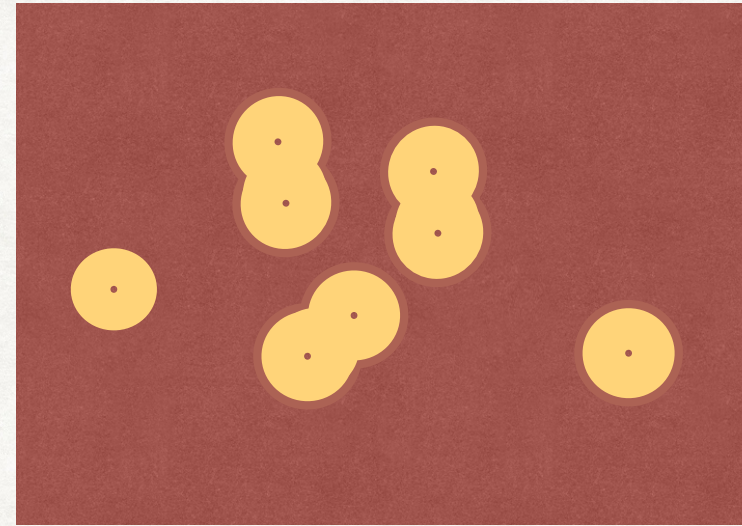


FIG. 1: Top left panel: example of a simulated 12-square-degree convergence map in the fiducial cosmology, with intrinsic ellipticity noise from source galaxies and $\theta_2 = 1$ arcmin Gaussian smoothing. A source galaxy density of $n_{gal} = 15/\text{arcmin}^2$ and $z_s = 2$ was assumed. Other three 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 $\kappa = 0.1$ (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.

ジェット内の粒子分布への応用

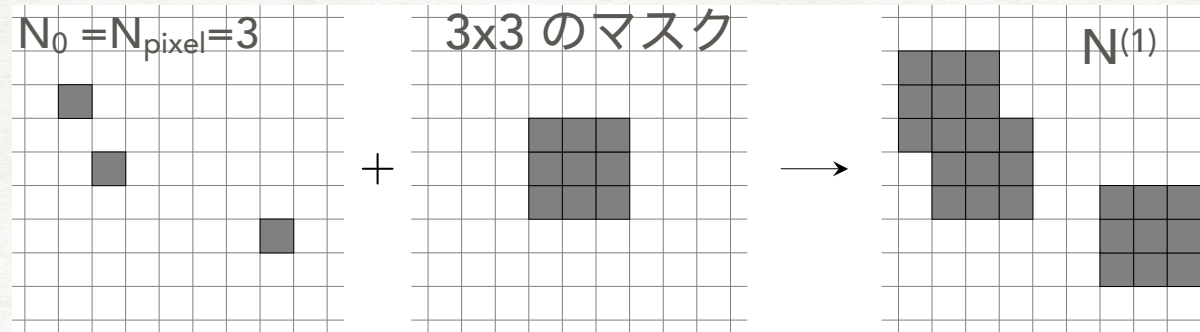
- 基本的に良いことばかり

- それぞれの点の持つ、周囲の点との間の距離の情報を、同等に取り扱う。ジェット発見のアルゴリズムと同じコンセプト
- 重なったエリアを2重に数えないのでパーティンシャワーの同方向への分離に対して安定
- ジェットアルゴリズムの特徴量(Jet Area)と同じ量
- 全ての粒子間距離をまとめて取り扱うので、統計的なふらつきは少ない。損失関数の最小化と相性がいい。
 - CNNアルゴリズムは、ジェットイメージの一つのピクセルごとの揺らぎを直接取り扱わないといけない。



ジェットイメージに対する実装

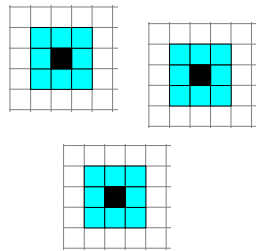
- ジェットイメージのピクセルに対して 3×3 , 5×5 ,のマスクを準備する



- N_1/N_0 が持つ情報の例

バラバラの点

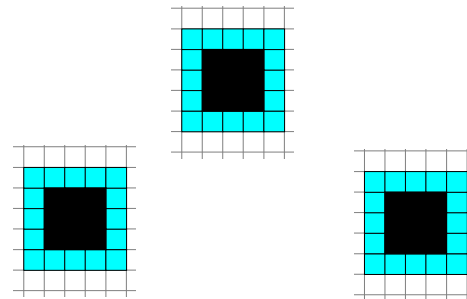
$$N_1/N_0 = 9$$



VS

all pixel appear in 3x3

$$N_1/N_0 = 16/9 = 1.78$$



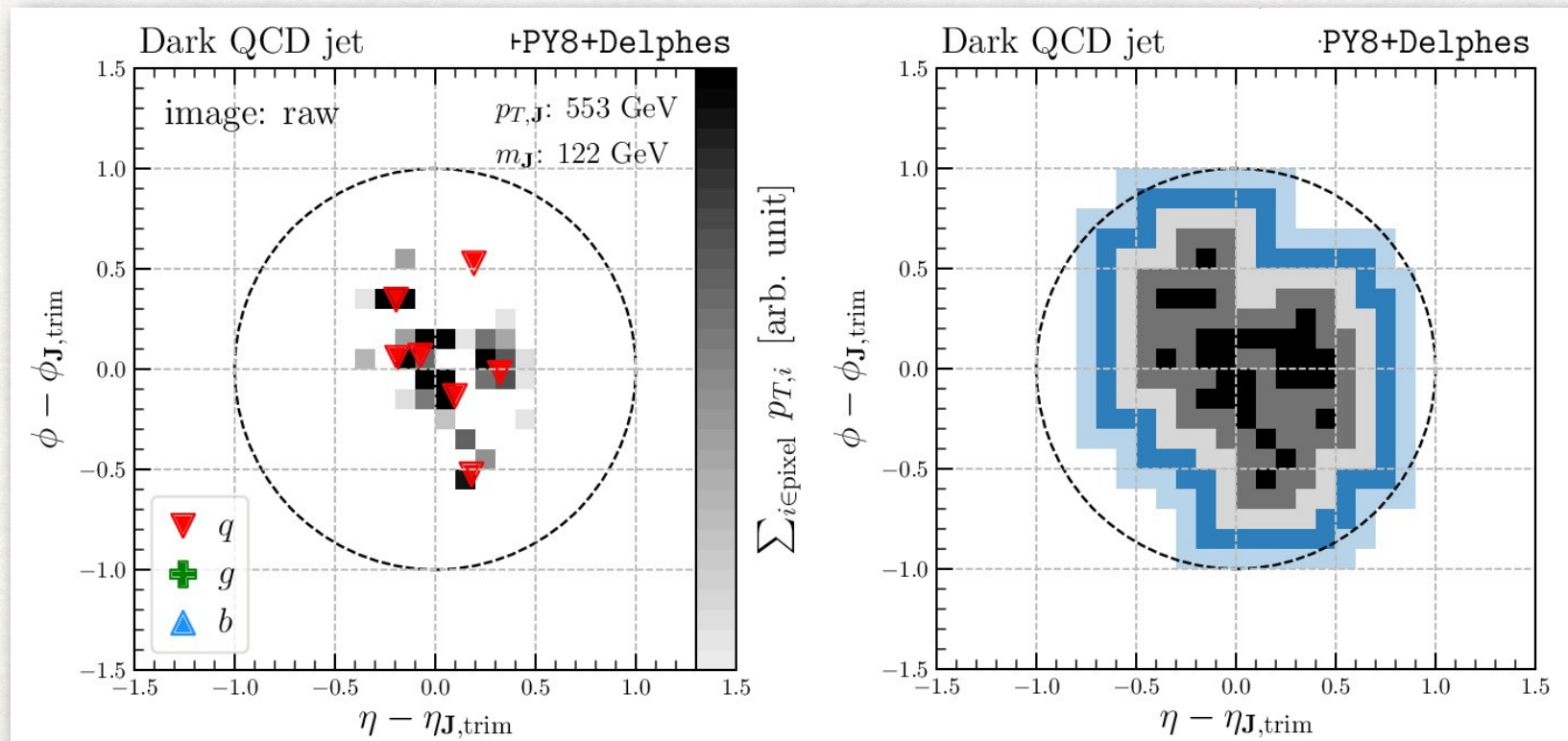
ミンコフスキー関数を使った ジェット分類

DARK JET の場合

Lim, Nojiri in preparation

- Dark Jet $pp \rightarrow Z' \rightarrow q_D q_D \rightarrow$ dark Parton shower $\rightarrow \rho_{\text{diag}} \rightarrow qq$
- 粒子がたくさんあるが、カラーシングレットなクラスターになっている状態

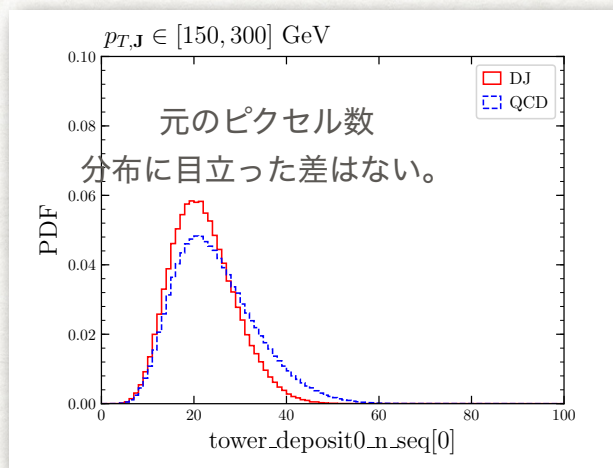
$m_\rho = 20 \text{ GeV}$



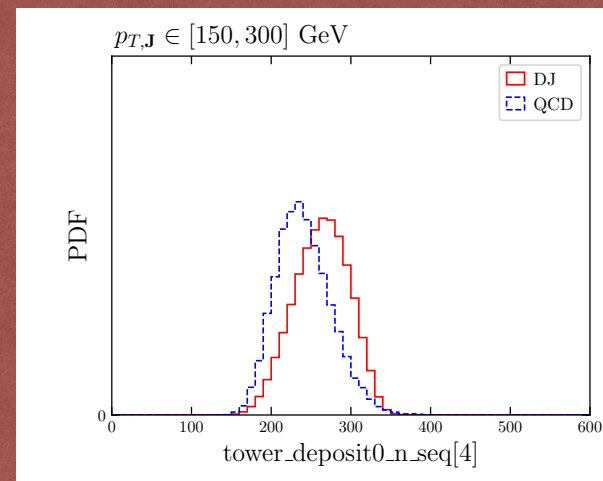
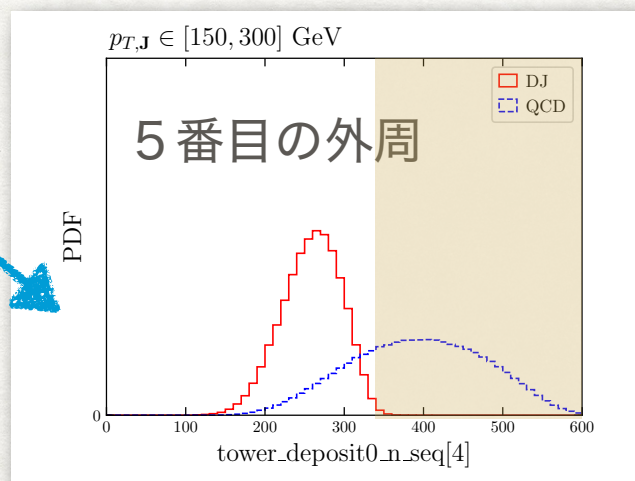
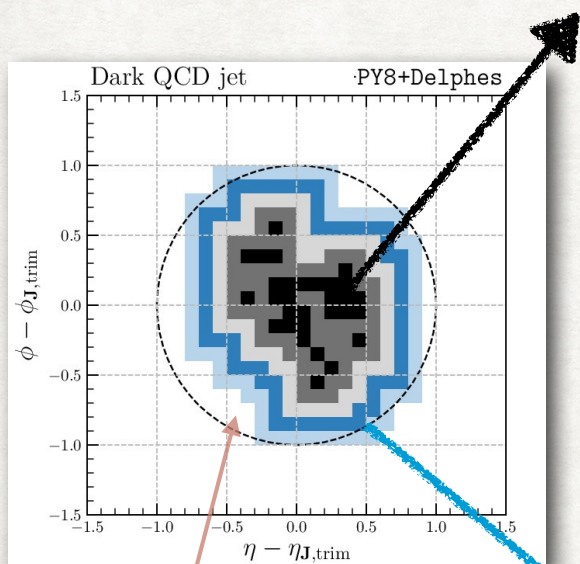
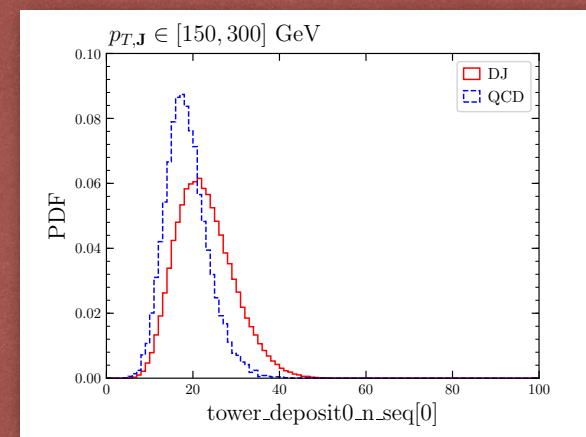
CNN の学習結果

- $m_p = 20\text{GeV}$, $300\text{GeV} < p_T < 400\text{GeV}$ CNN のイベント選択は、MF($k > 2$) の分布でカットをかけていた。

元の分布



CNN でカット



何もない空間に注目している。

CNN とミンコフスキー関数

- 畳み込みニューラルネット (CNN) : QCD と t, Z, W , 新粒子のジェットイメージの MF が十分に違っていれば、CNN は MF を学習しているだろう。

実際 MF は 2×2 のフィルターへのヒットを以下の数値を与えることで計算できる

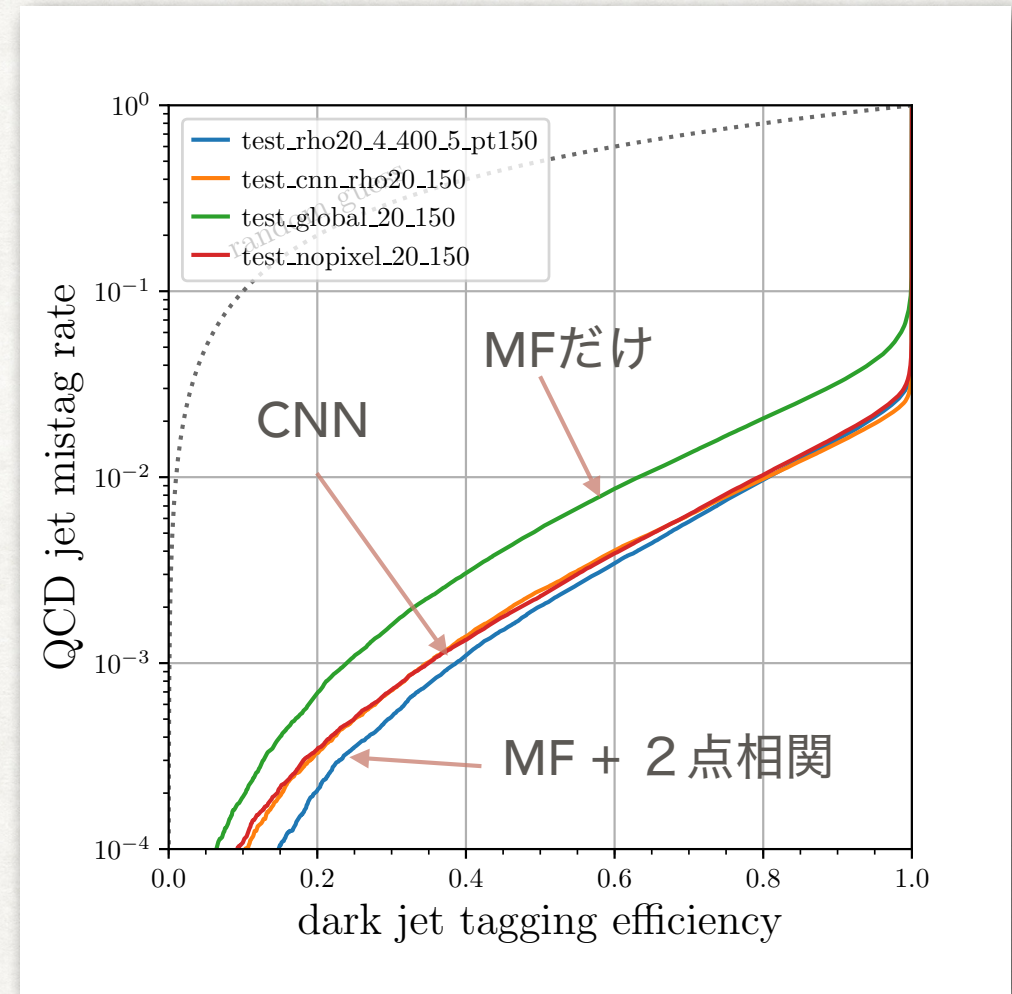
$$v(f_i) \rightarrow (A_i, P_i, \chi_i), (A, P, \chi) = \sum_i v(f_i)$$

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

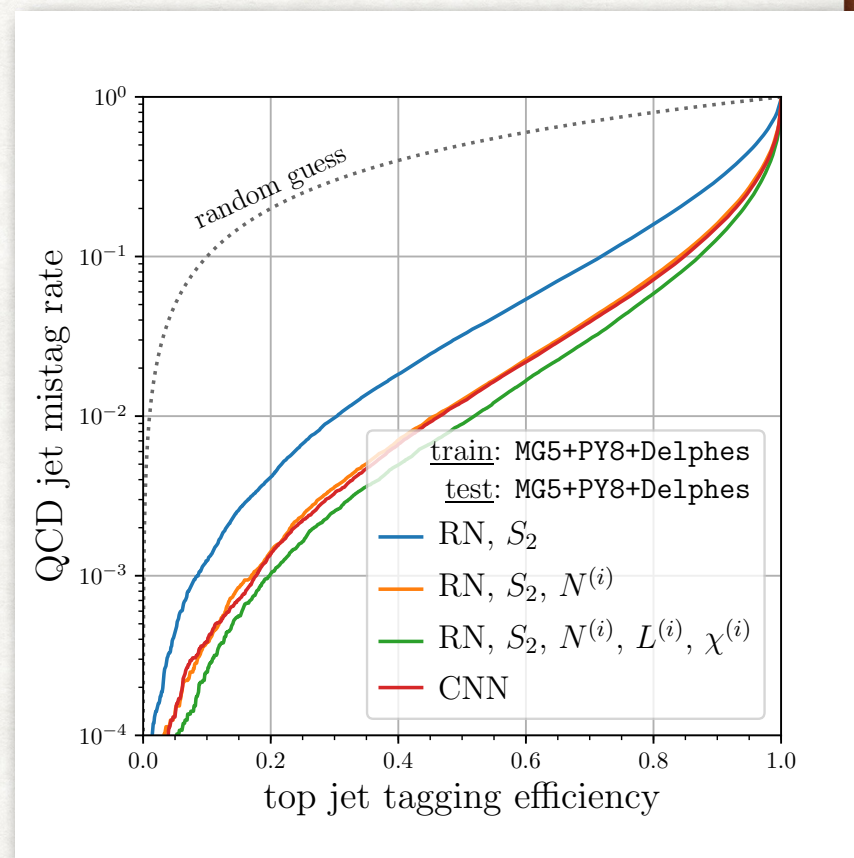
もっといいNN を作ろう!

- RN(Relational network) :
- Jet mass, jet pT
- $N^{(0)} \sim N^{(3)}$ [pT cut なし、2 GeV, 4 GeV]
- エネルギーに依存する情報
 - S2 : C correlator(Energy correlator) $f(\theta) = E_i E_j \delta(\theta - \theta_{ij})$
 - Tkachov (hep-ph 960138) Lim, Nojiri
1807.03312, Chakraborty, Lim Nojiri
1904.02092 ~ τ の任意の β の情報を担っている。
- Top 等の場合、3点も重要はず ;
 - [Leading subjetの粒子との2点関数]
 - [Leading subject の粒子を除いた2点関数]



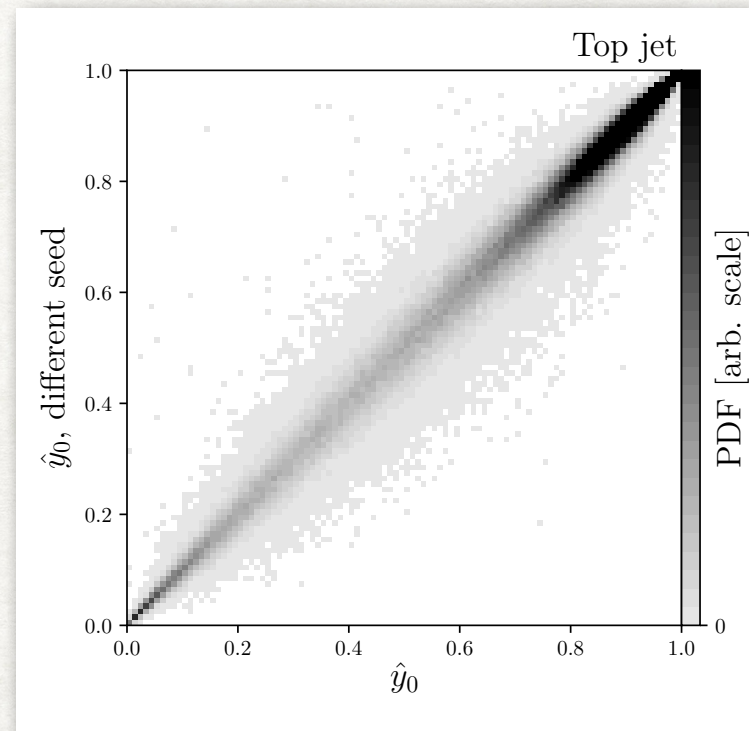
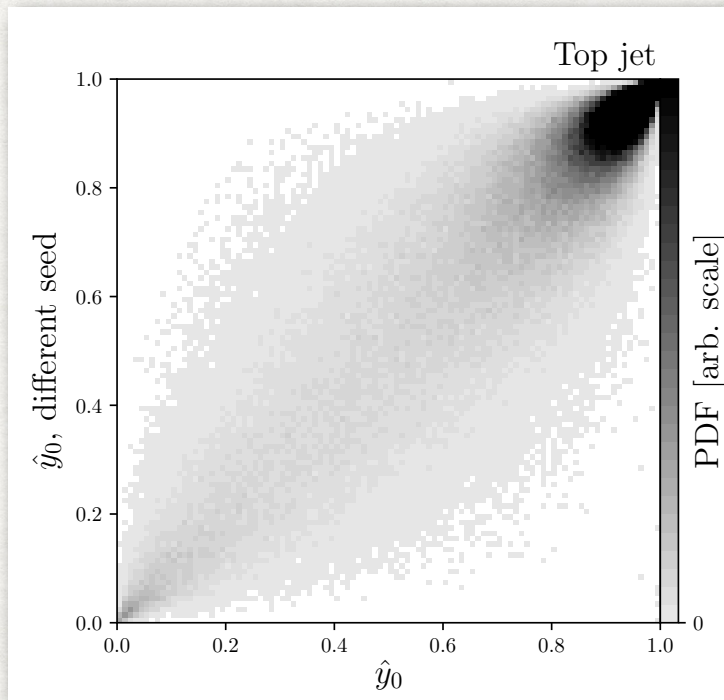
TOP VS QCD ジェット

- pT 500GeV~ 600 GeV までのtop jet.
- Area 情報だけ使用した段階ではCNNとRN は同等
- RN に図形の面積だけでなく、周囲の長さ、オイラー標数も情報として追加することで、CNN を凌駕
- Dark Jet の場合と違って、Area 分布が劇的に違わなかったため、CNNはMFを特徴量として認識できなかった。



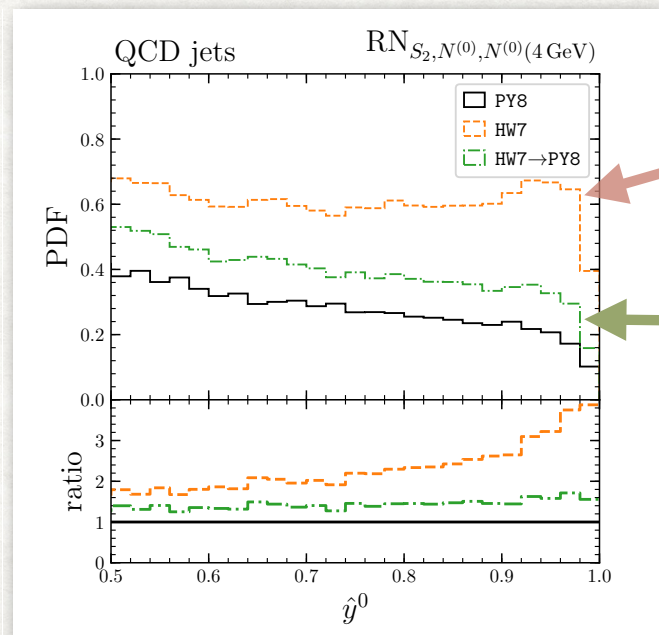
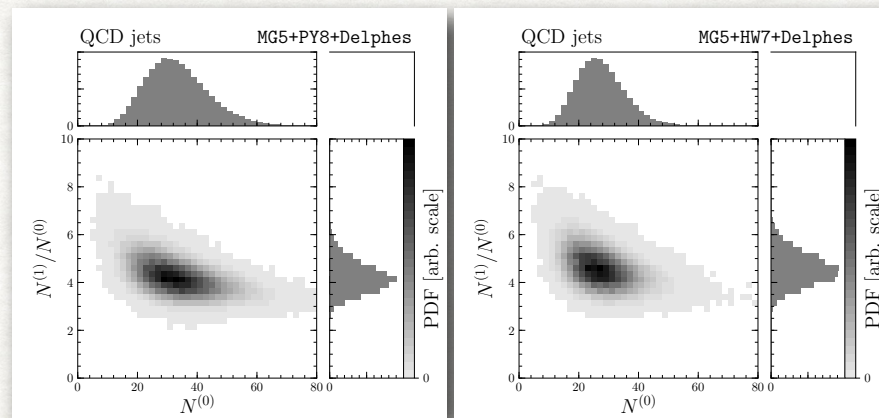
LOCAL MINIMUM 問題の改善

- CNN のloss function の最小化で、「真のminimum にたどり着く」ことはあまりない。
- ROC は安定しているが、個々のイベントに対して、違うseed で使ったclassifier は違う結果を出す。
- RN + MF は input が少ないのでevent ごとの結果も遙かに安定 (900-> 85)



MC TURNING とかCALIBRATION とか

- TQCD ジェットは event generator が違うと結構違う分布を予言する。
- 特に、QCDジェットの粒子数と広がりが大きく違う、しかも実験データとも合わない。(モデルの改善が必要)
- [MC を"実データ"で補正] MF の値が同じになるように、イベントにウエイトをつけると、一致がよくなる。



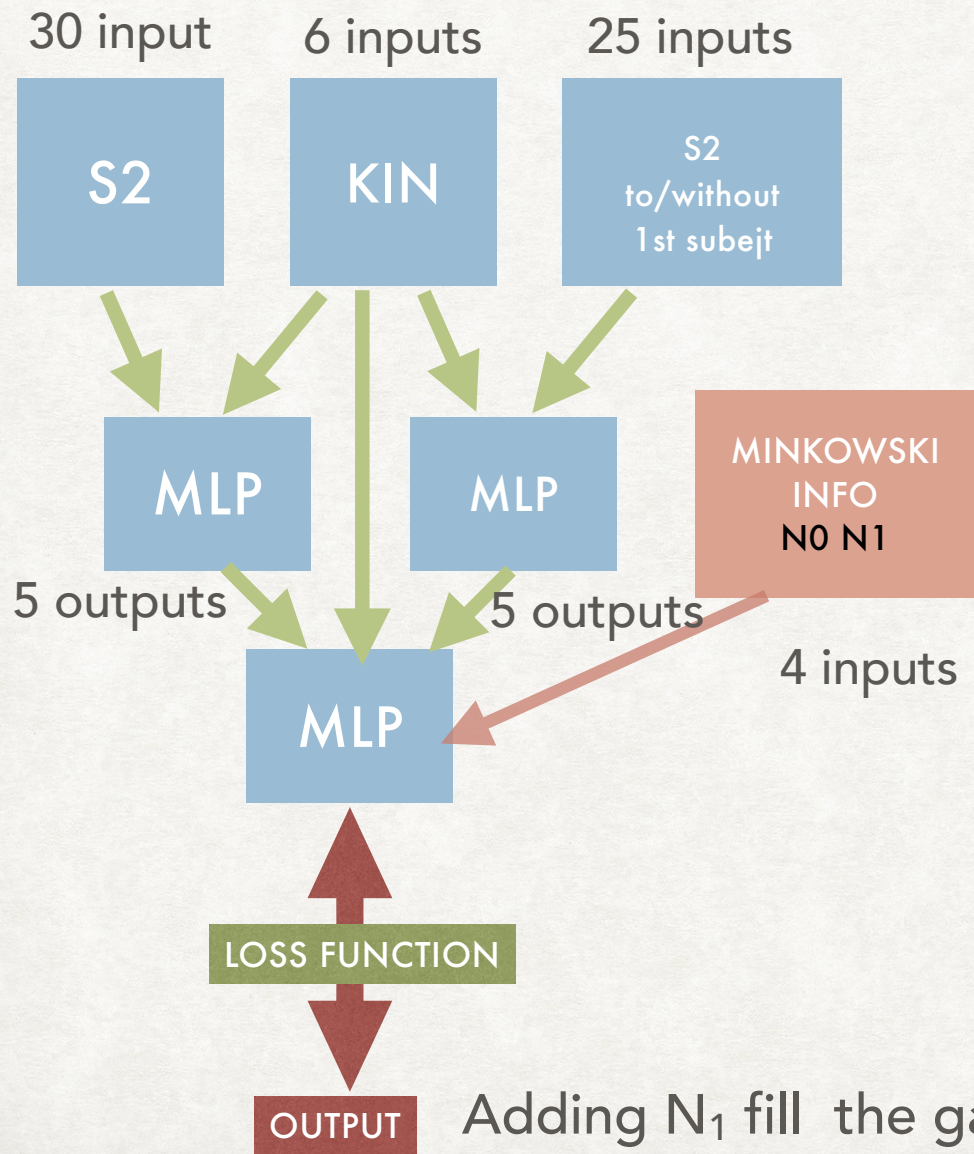
Phythia で training した classifier で
Herwig のQCD jet を分類

N0 分布を補正した後

深層学習と物理

- MF: カラー構造の違う粒子の比較的運動量が小さいエネルギーdepositの分布を効率的に捉えているように見える。
- Dark Jet vs QCD 深層学習が自律的にミンコフスキー汎関数（物理）を見つけてきた。
- Top jet vs QCD 深層学習は、支配的な特徴量(subjet 分布) しか見つけたれなかった。
 - もっとも特徴的な量に損失関数が支配されて、それ以上改善しない。
- pixelごとの Order(1) でふらつく情報ではなく、MF のような「まとめ指標」の方がキャリブレーションにも便利かもしれない
- 他の問題（例えば、displace vertex の発見など深層学習が得意な分野）にも使えるかも

おまけ NN のシステム



process

pp \rightarrow tt vs pp \rightarrow 2j

$500\text{GeV} < p_T < 600\text{GeV}$

$150\text{GeV} < m_j < 200\text{GeV}$

case 1

modulation for two point correlation

two point correlation + Kin \rightarrow 5 outputs

correlation to/without leading jet

\rightarrow 5 outputs

\rightarrow ROC

case2 + N0(number of active pixel)

\rightarrow ROC

case 3 + N0,N1 \rightarrow ROC

Adding N_1 fill the gap between CNN and our approach.