

b-tagの変数を調べる

2, Nov, 2012, 加茂直之

基本情報まとめ

- **jobOption**:MC11.105850.WH120bb_pythia.py
- **D3PD**:NTUP_JETMET
- **イベント数**:48990
(本当は5万、D3PD作るときにかなりミスしたみたい。)
- **JET数**:161951
- **Jet_Pt>10GeV?**
(ヒストグラムの見た目で判断。
QcdD3PDMakerにはminPt=7GeVという記述があった。別のところでカットかけているのだろうか)

b-tagをする各変数の検証

- 目標: Jetのb-tagを表す変数が3つあり、その変数の意味するところを調べる。
- そのうえで、どれが良い変数か知りたい。

- 次から載せているヒストグラムは
- jet_AntiKt4TopoEM_flavor_weight_JetFitterTagNN
- jet_AntiKt4TopoEM_flavor_weight_JetFitterCOMBNN
- jet_AntiKt4TopoEM_flavor_weight_GbbNN

をそれぞれ

jet_AntiKt4TopoEM_flavor_truth_label

=0 (それ以外のjet)

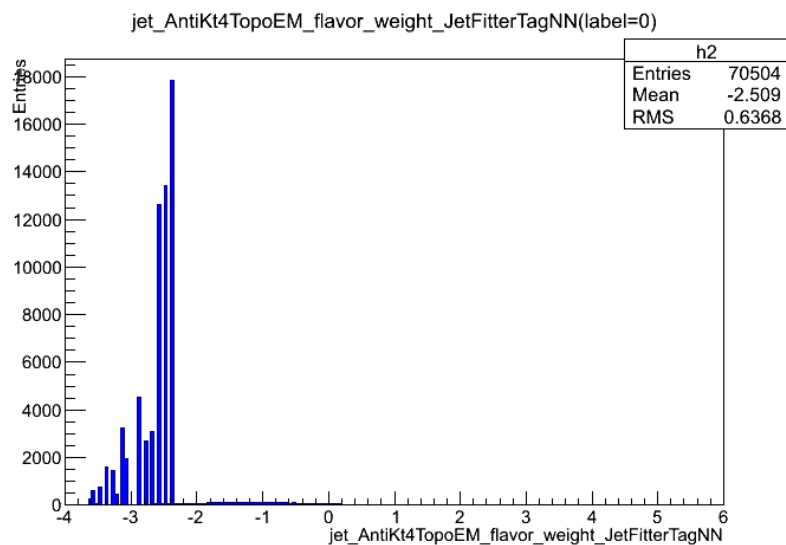
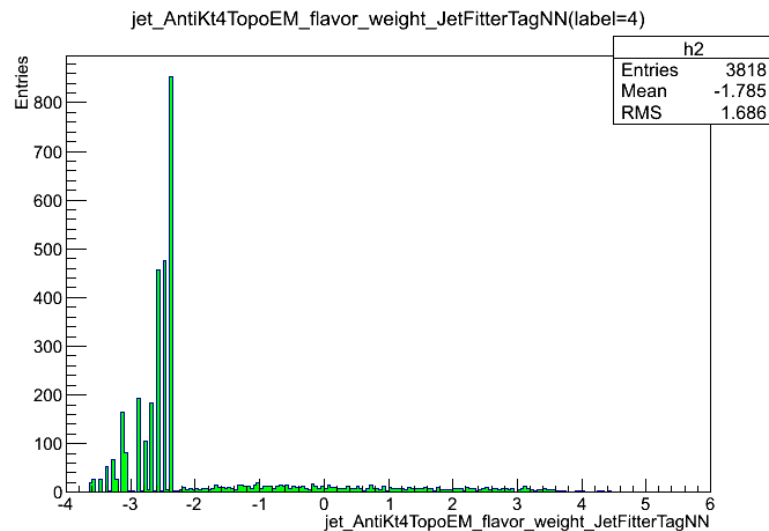
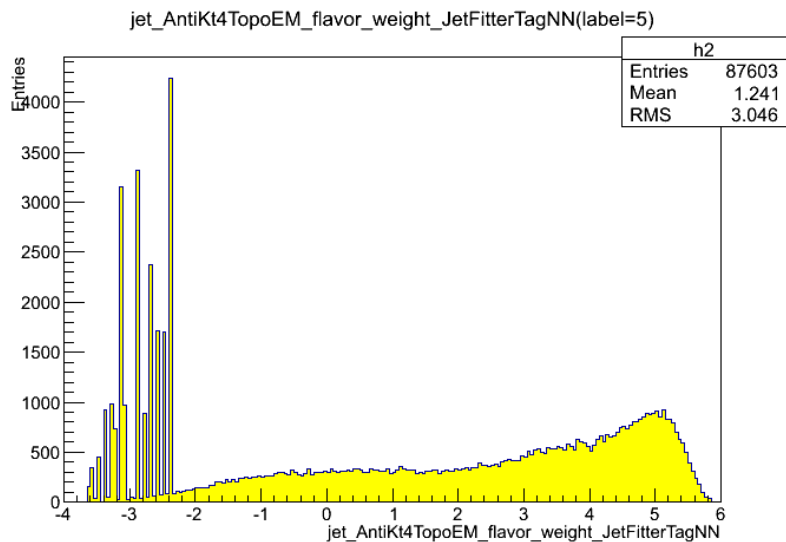
=4 (charm)

=5 (bottom)

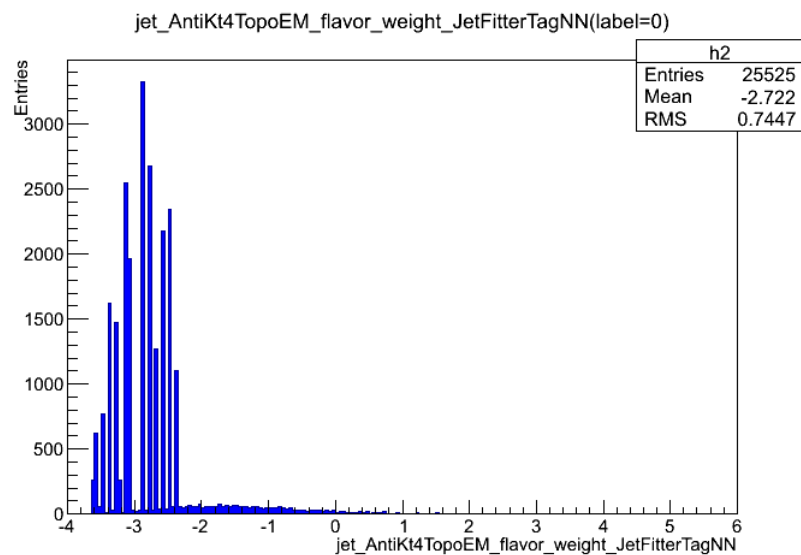
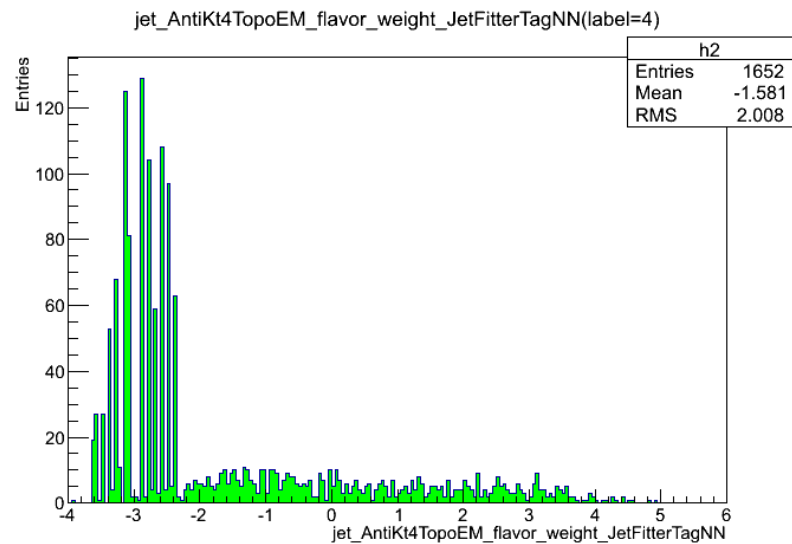
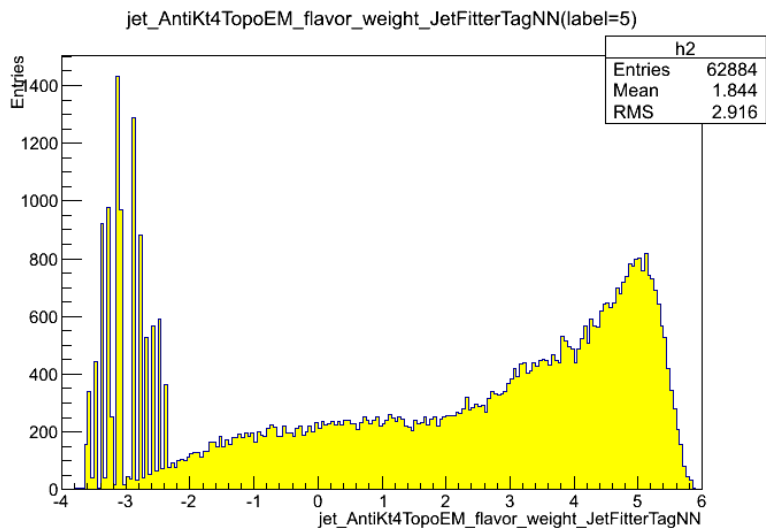
、カットありなしで場合わけしたもの。

カットは $Pt > 20 \text{ GeV}$, $-2 < \eta < 2$

カットなし、JetFitterTagNN(黄:bottom,緑:charm,青:その他)



カットあり、JetFitterTagNN(黄:b, 緑:c, 青:他)

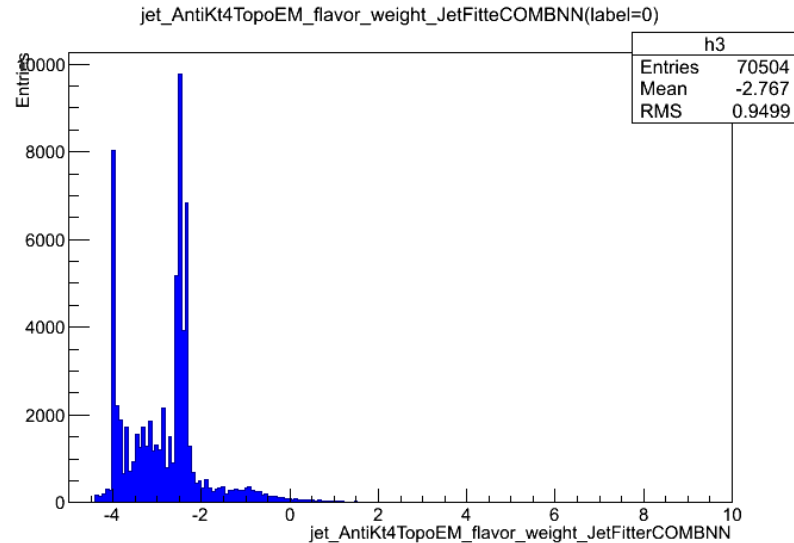
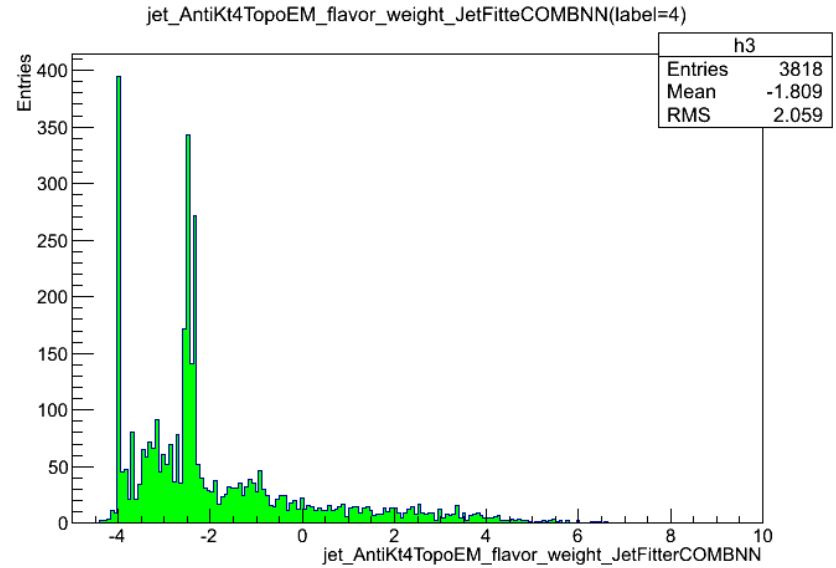
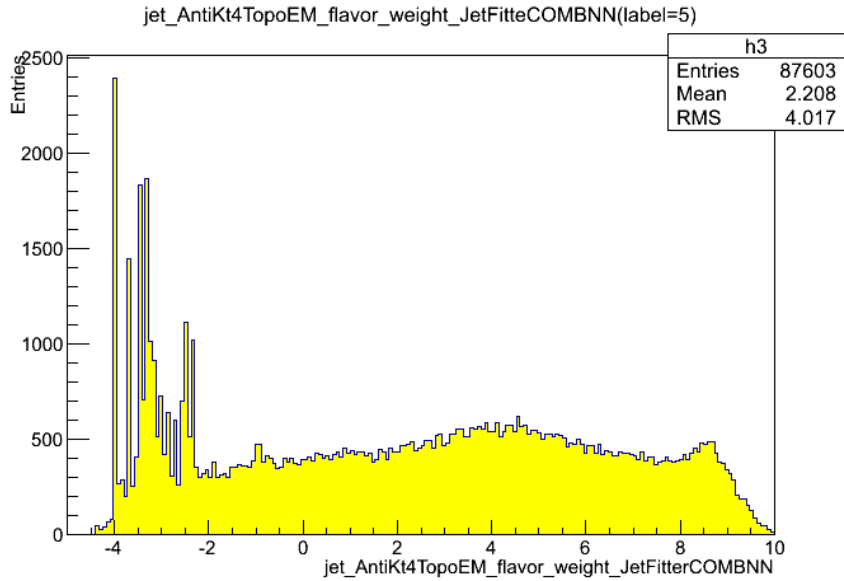


TagNNの結果、考察

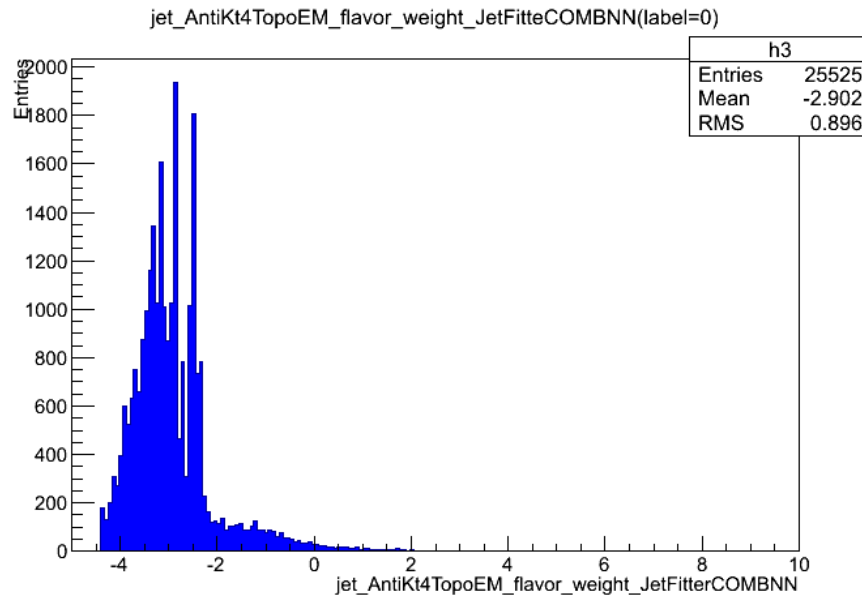
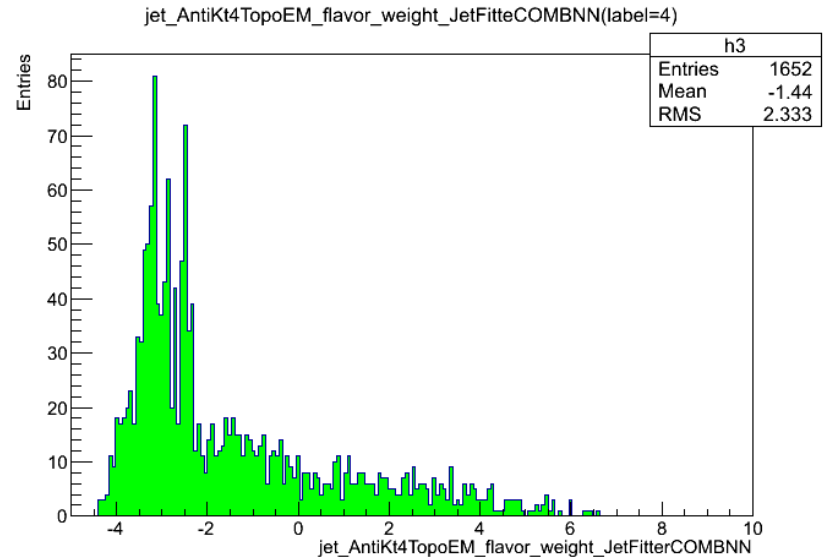
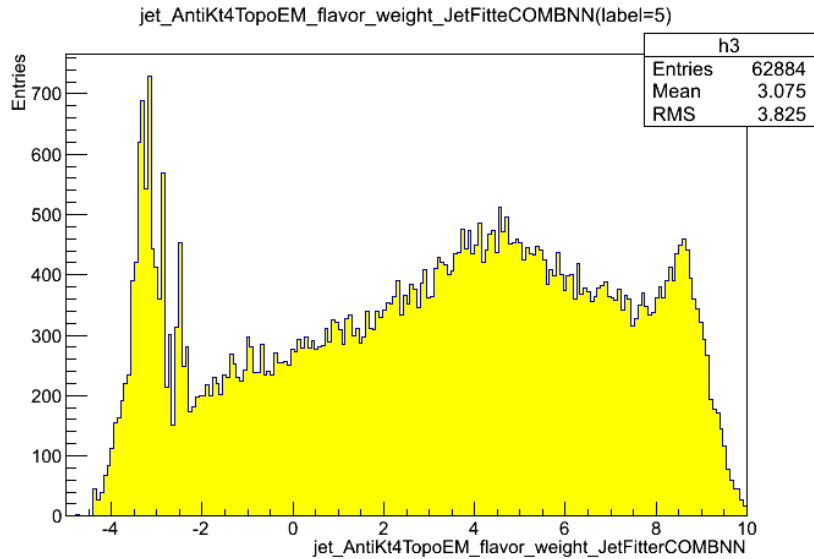
b-jet(label=5のjetをb-jetと書きます)

- b-jet以外のjetはだいたい-2以下になっている。
 - b-jetは全領域に広がっている。
- この変数でb-jetを定義するならTagNN>-2にする？
- b-jetは5付近でピークがたってるが、これは意味があるのか？
- このあたりにあるjetをなにか調べてみる
- カットかけたら、値が小さい領域のEntryが少し減った
 - (他の変数についても同じだけど)c-jetが数値が大きいところでも、Entryが少しあるが、これはbがcに崩壊したときに、b-taggingをミスしたもの？

カットなし、JetFitterCOMBNN(黄:b, 緑:c, 青:他)



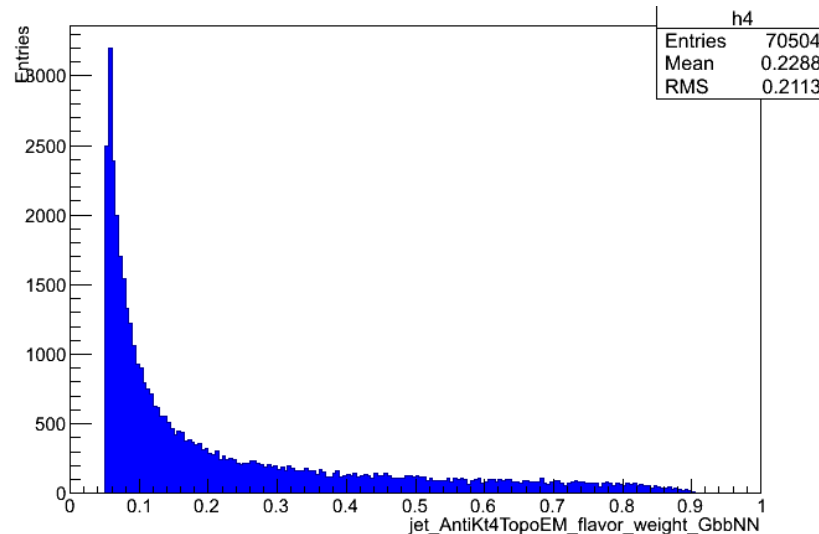
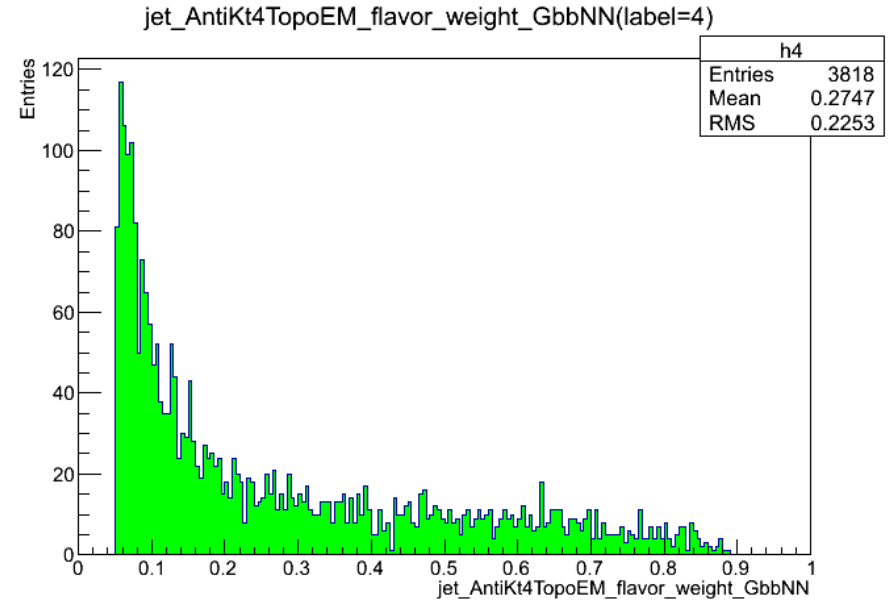
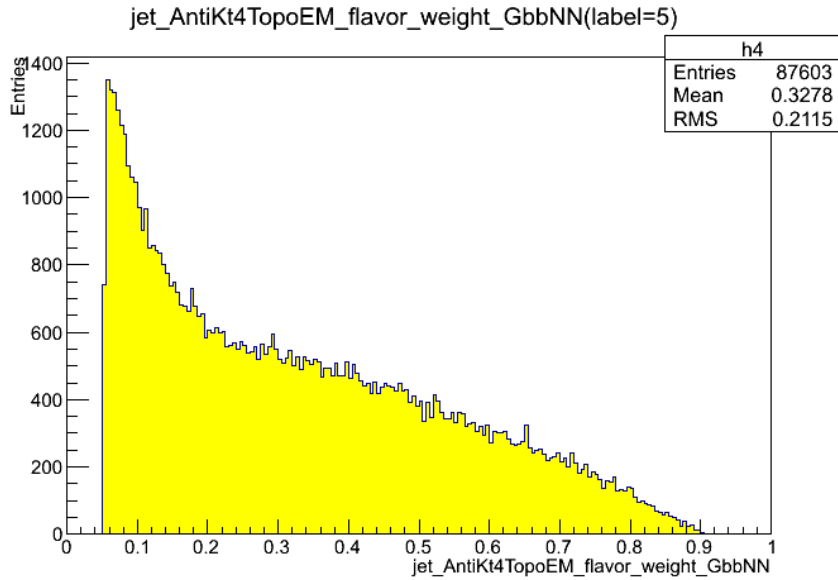
カットあり、JetFitterCOMBNN(黄:b, 緑:c, 青:他)



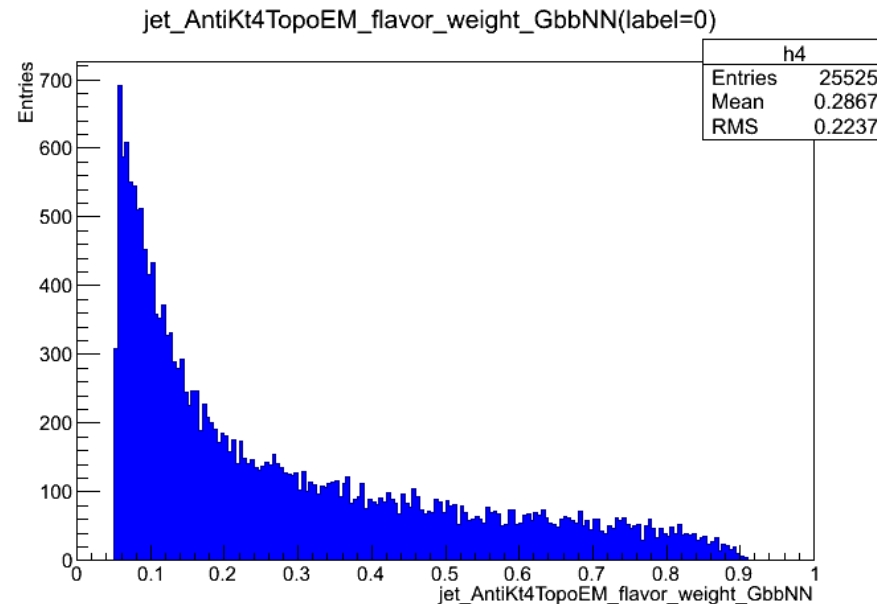
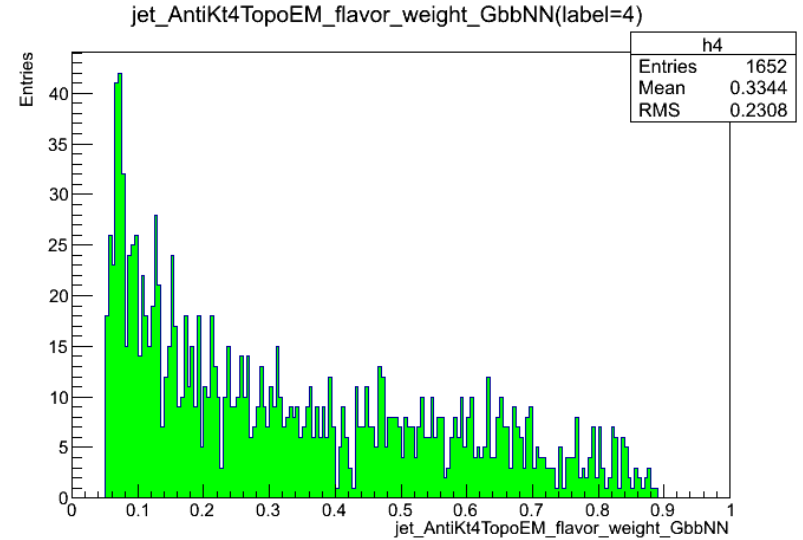
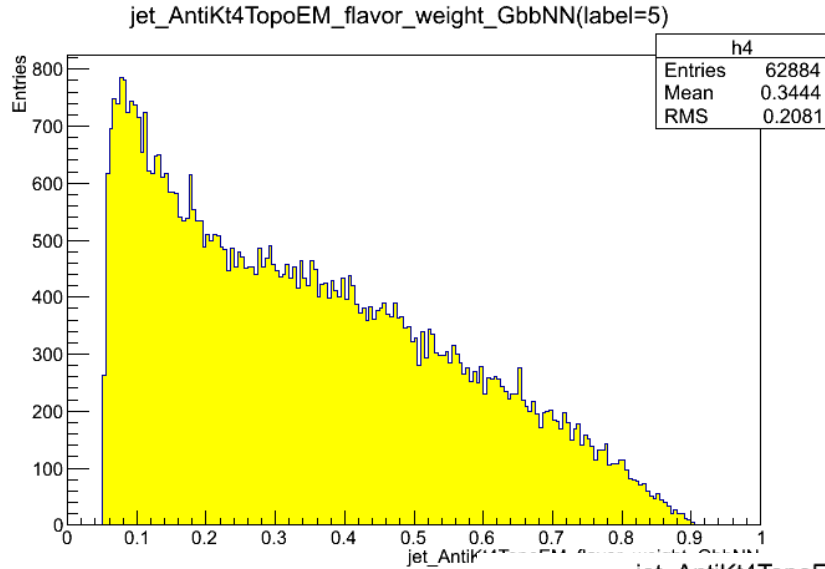
COMBNNの結果、考察

- こちらも-2という数字がbとそれ以外をわけているように見える。
- ただしTagNNと比べb-jetでもにかなりEntryが-2以下に広がっているし、
- 他のjetも-2以上に広がっている(見た目)
- B-jetのピークが4と9あたりで立っている。
→TagNNのピークと比較してみる？

カットなし、GbbNN(黄:b, 緑:c, 青:他)



カットあり、GbbNN(黄:b, 緑:c, 青:他)



GbbNNの結果、考察

- JetFitter*とすこし毛色が違う感じ。
 - 0.1か0.2ぐらいで他のjetが少なくなっているように見える。
 - ただ、どのjetも0.1以下でEntryが増えているし、全領域にEntryが広がっている。
- 今のところ、他のと比べ、b-taggingには使いにくそう

まとめ

- 数値の大きさを”bらしさ”を定義してるいるというのはある程度正しそう。
 - ただし、上限があるし、数値が大きいほどb-jetが多いわけでもないので、変数の意味はまだ調べないとわからない。
 - 見た目で判断するとJetFitterTagNNが良さそう。
 - (flavor_truth_labelを今は完全に信じるとして)TagNNを使うと、b-jet以外をほとんど切ることはできるけど、b-jetも一緒に結構切ってしまうている。
 - この変数がどこまでの情報を引き出せるかよくわからないが、もう少しb-jetを拾いたい。
- とにかく、まだ総合的に情報が足りないので
→ピークが立っている領域のjetを調べる
→変数同士の相関を見してみる、あるいは両方使ってb-jetを定義してみる
- という感じで考えている。