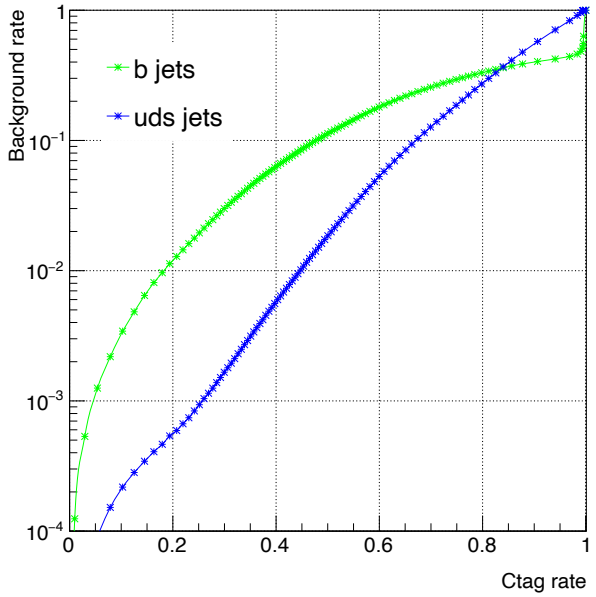
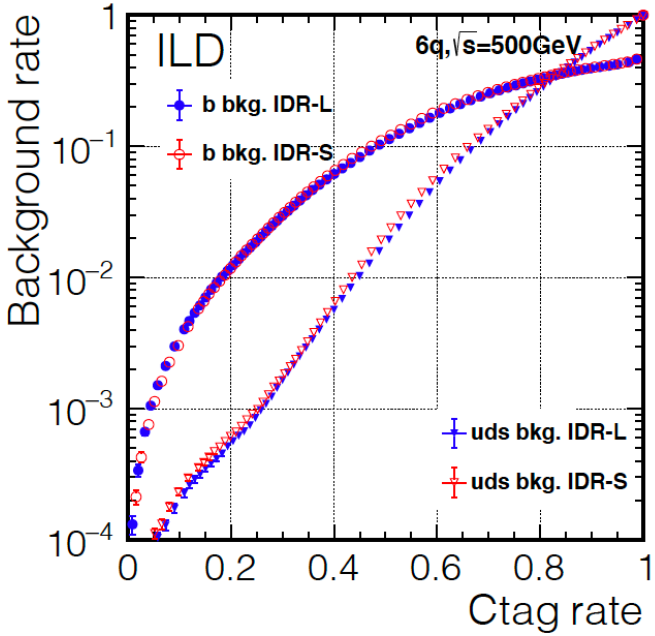
News

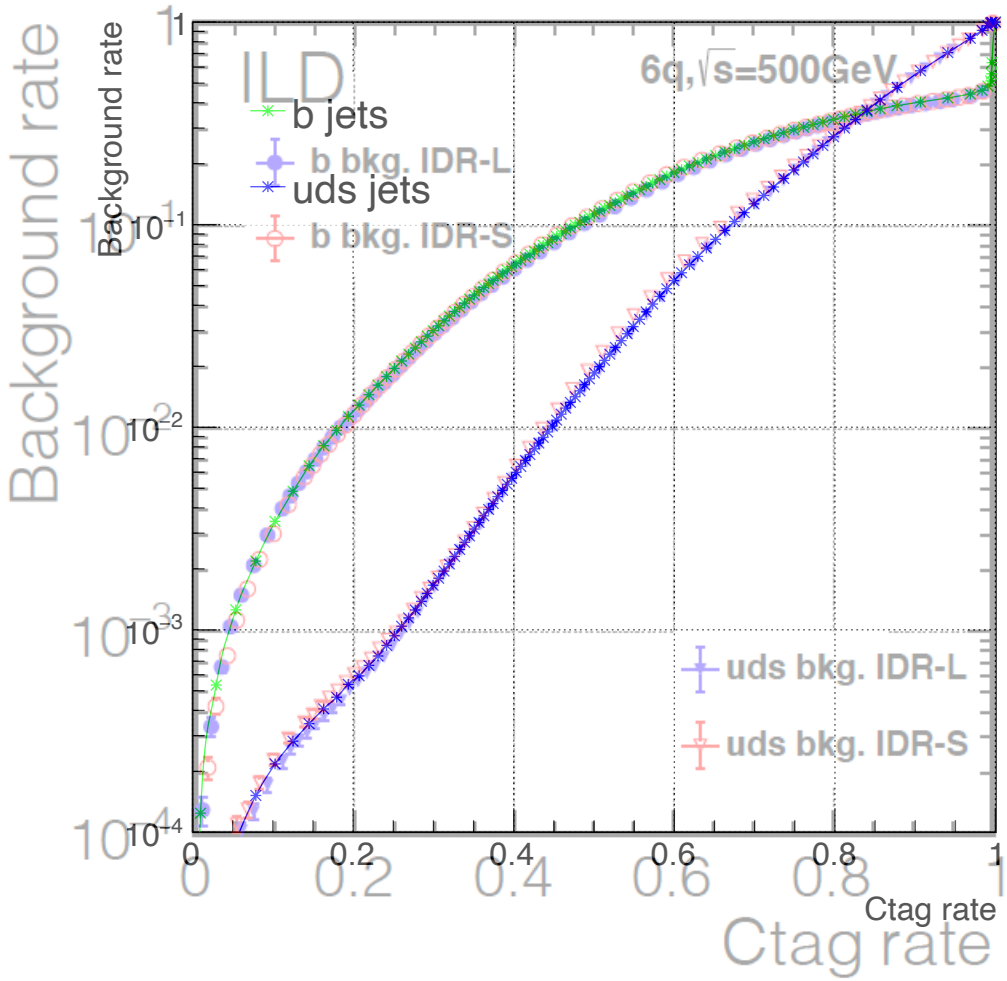
- re-produced LCFIPlus results
- solved memory problem on GPU
- run training on 4 GPUs in parallel (batch size increased to 200)
- learning rate is halved if validation loss stagnates for 10 epochs

- trainings:
  - with more constituents
  - with more variables

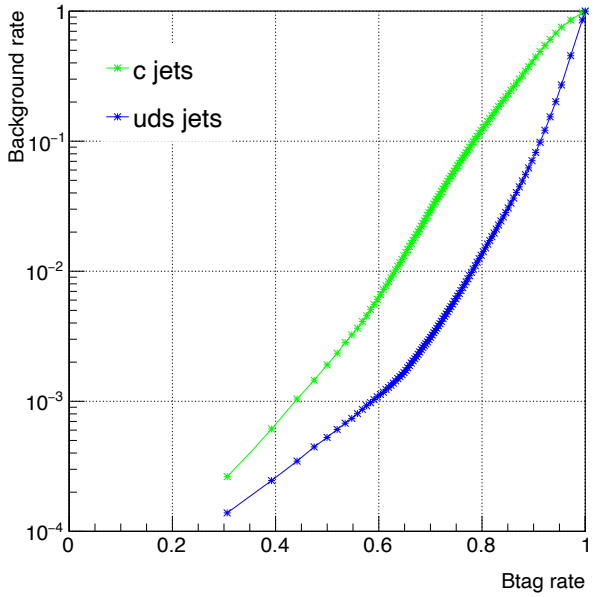
Performance LCFIPlus

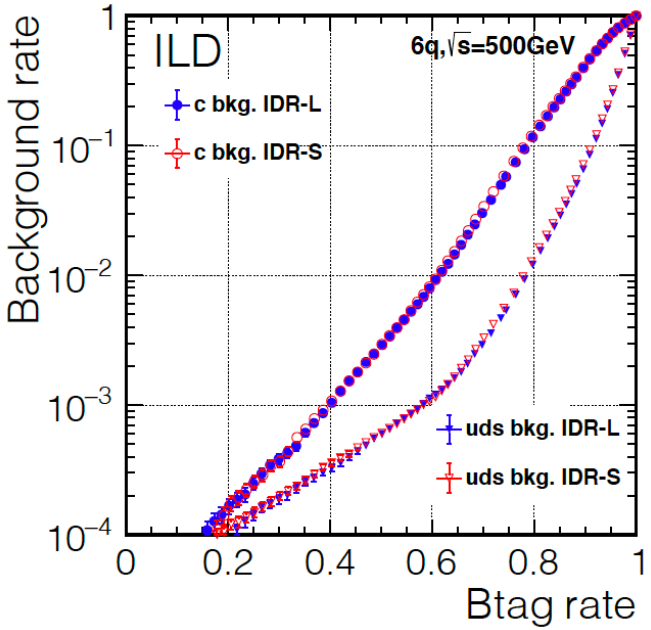


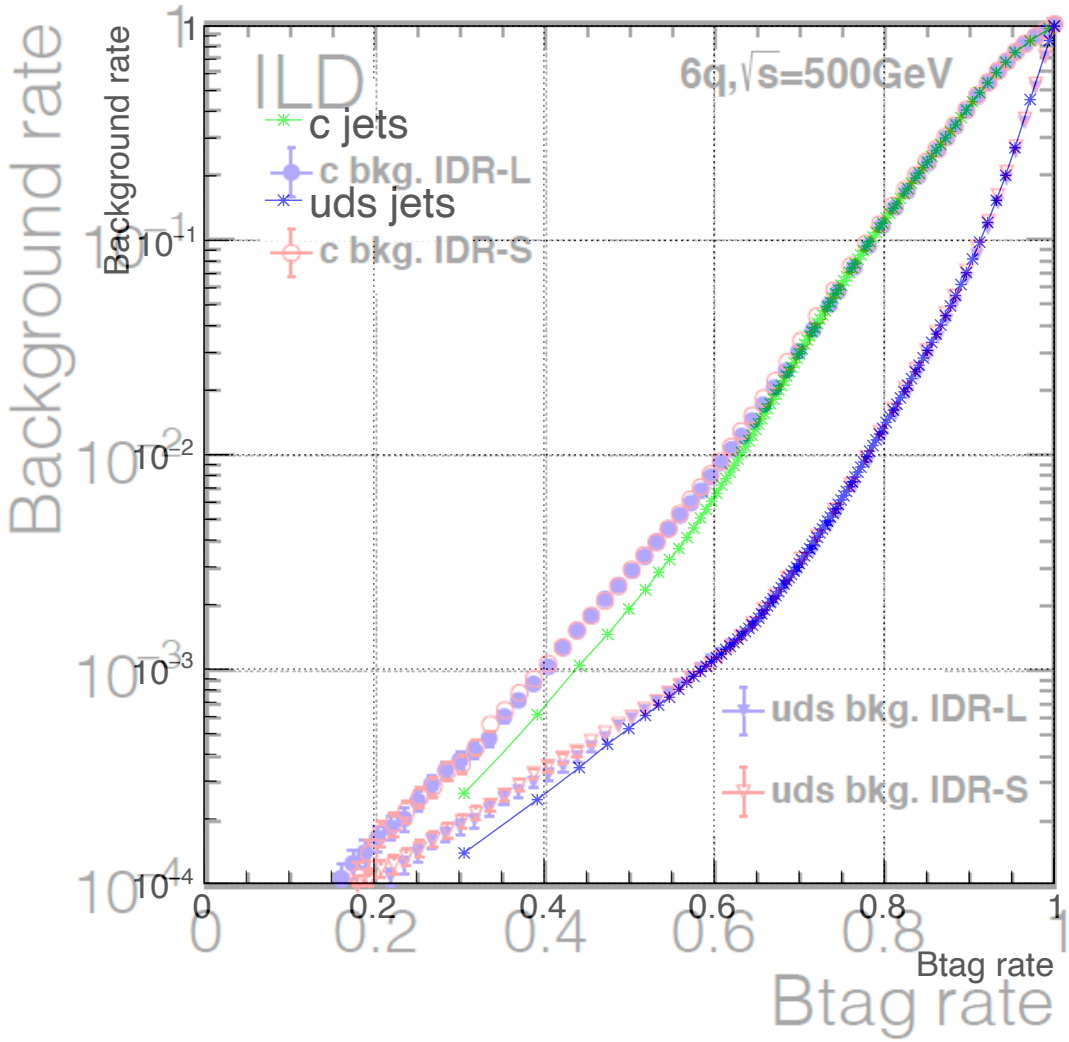




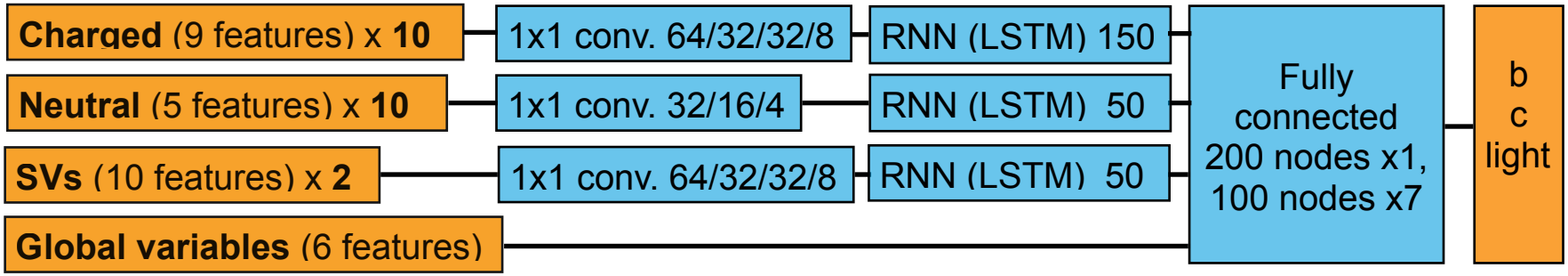
Performance LCFIPlus





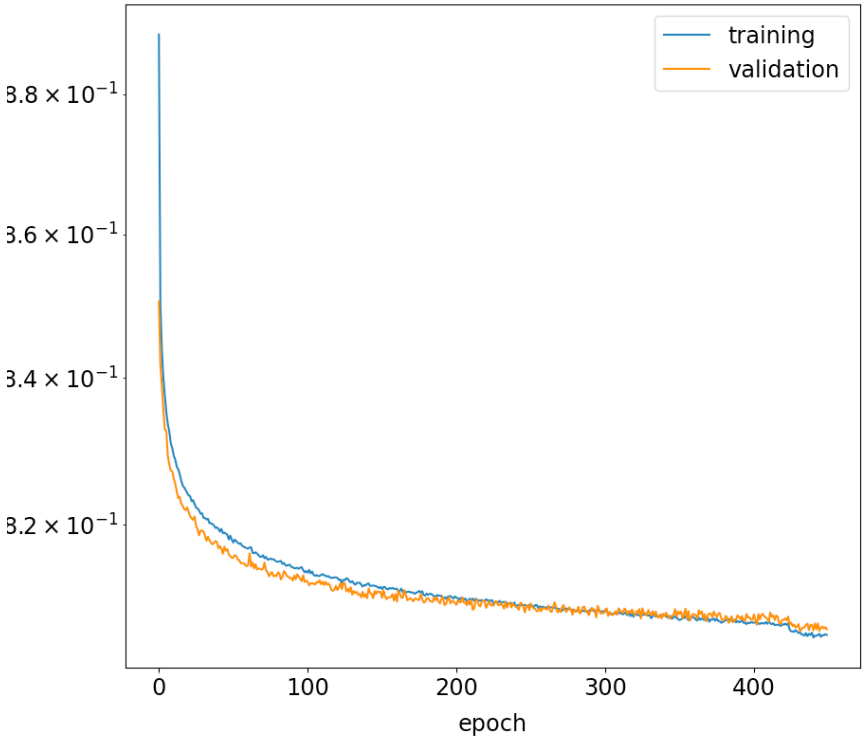


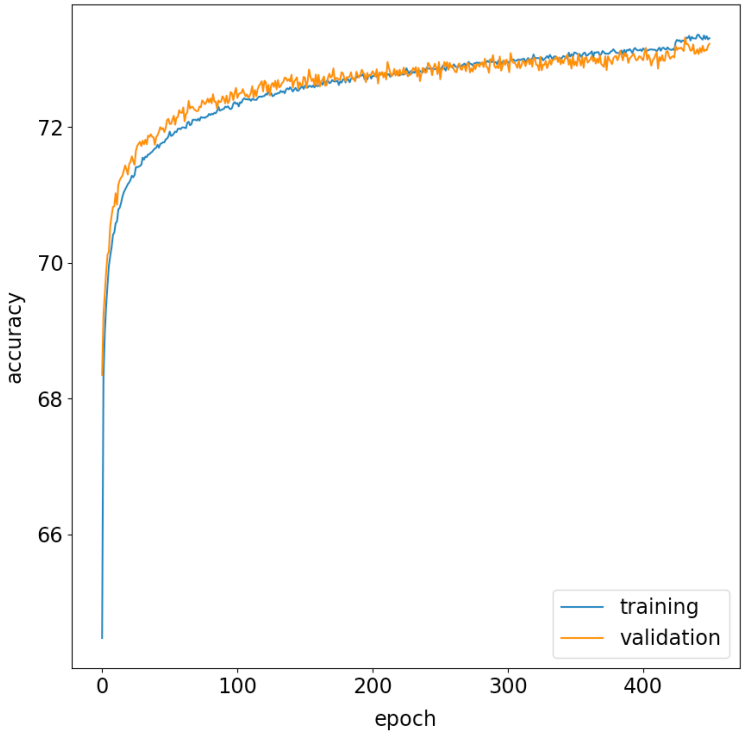
- reminder architecture:



- **new trainings:**
  - DeepJet architecture with more constituents
    - **Charged** (9 features) x **25**
    - **Neutral** (5 features) x **25**
    - **SVs** (10 features) x **2**
    - **Global variables** (6 features)
  - DeepJet architecture with more variables
    - **Charged** (23 features) x **10**
    - **Neutral** (6 features) x **10**
    - **SVs** (16 features) x **2**
    - **Global variables** (7 features)

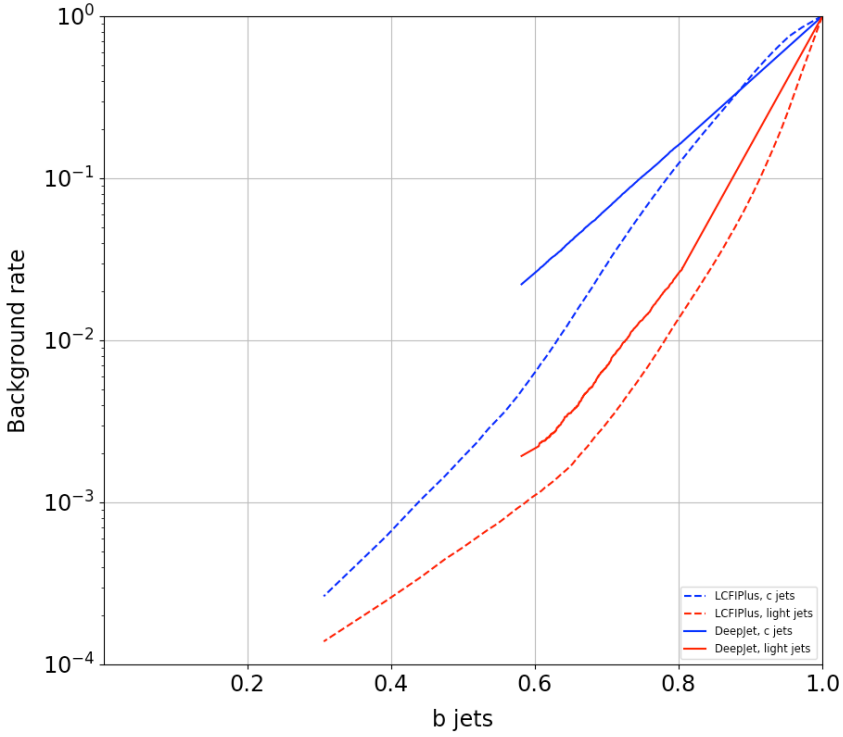
Results: training with more constituents

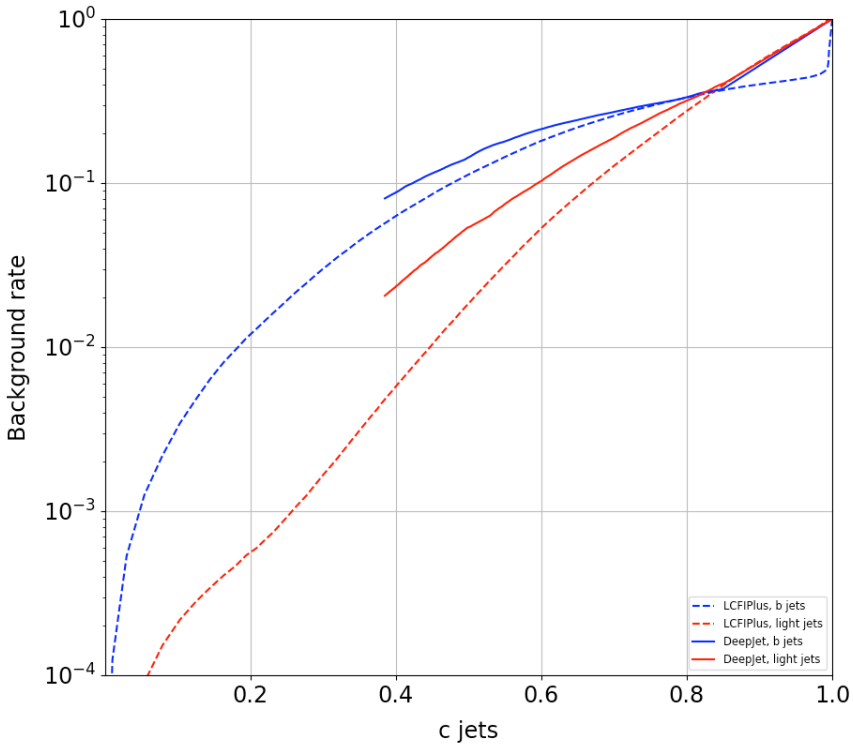




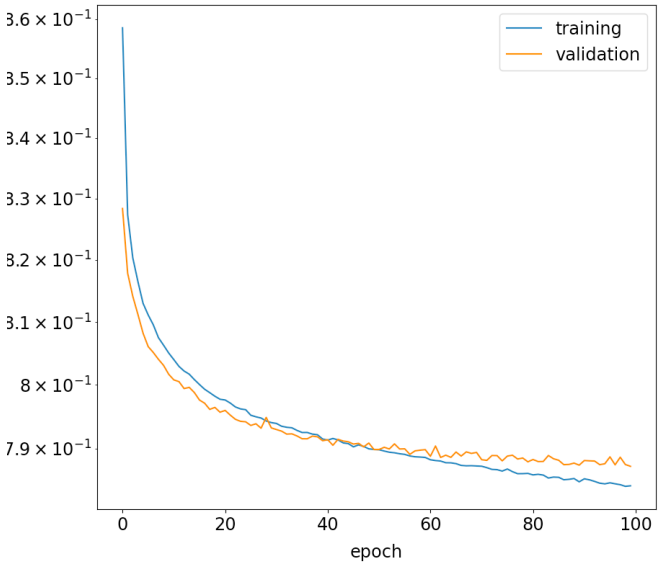
Results: training with more constituents

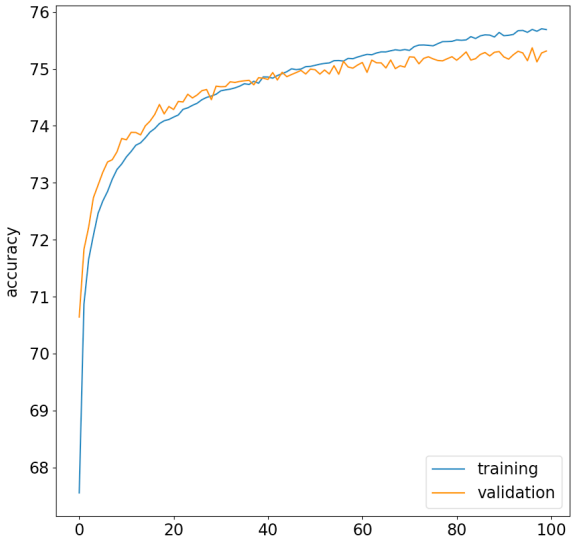
- ROC curves on validation sample



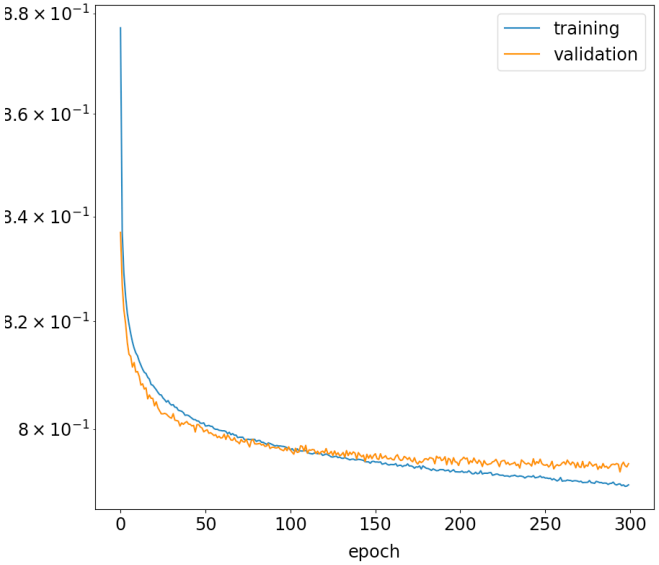


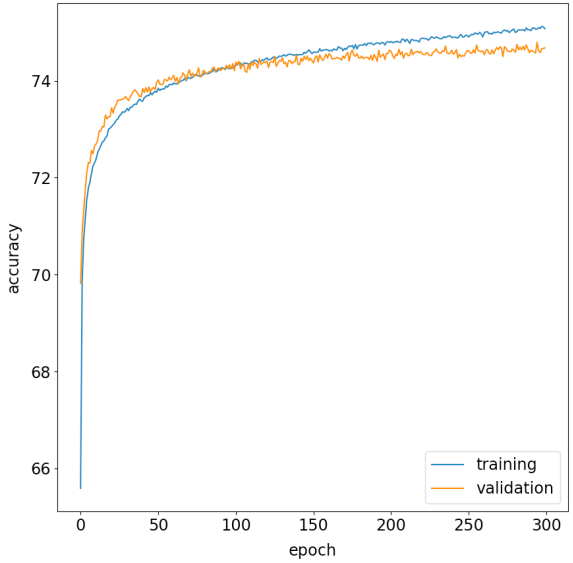
Results: training with more variables





dropout rate 0.1

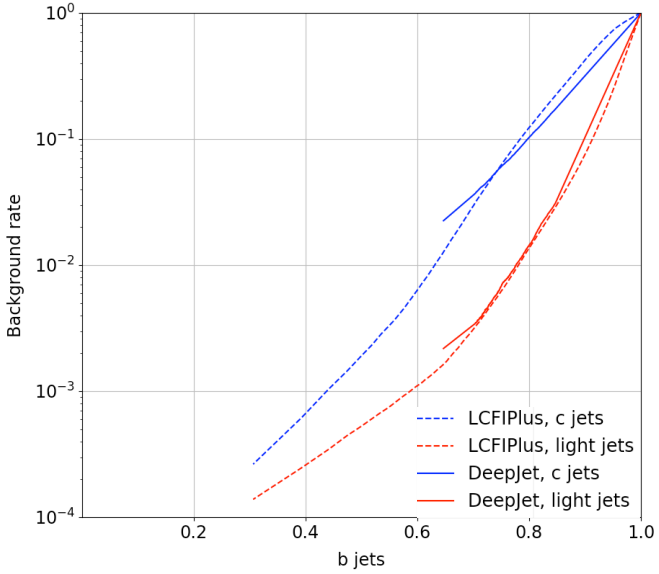


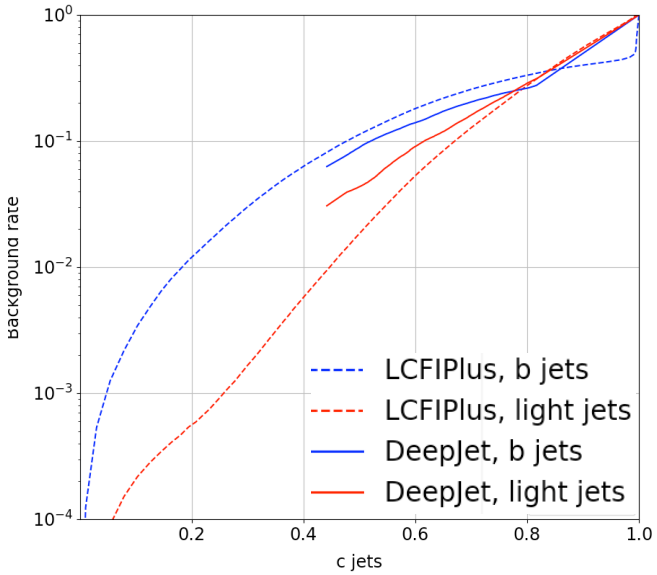


dropout rate 0.15

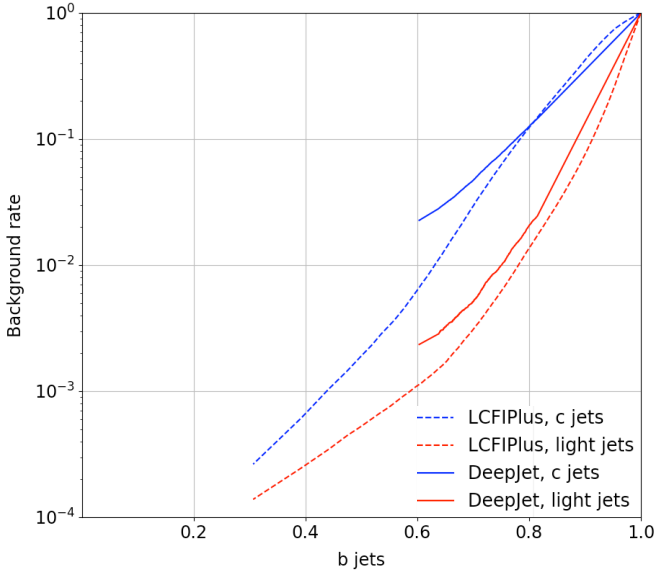
Results: training with more variables

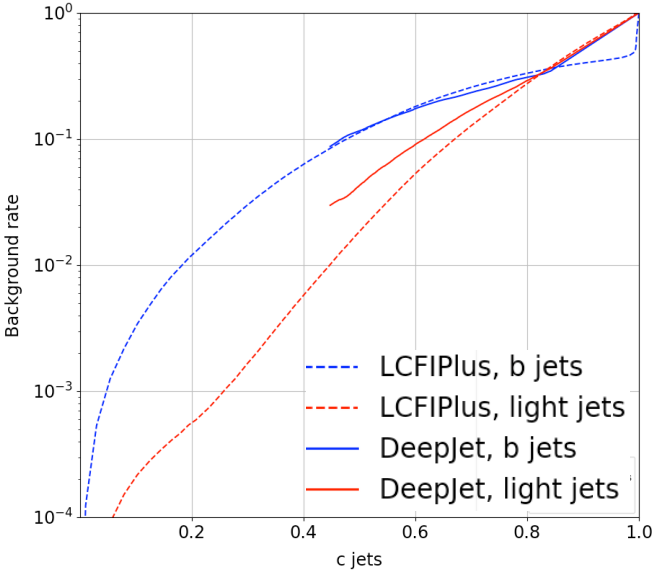
- ROC curves on validation sample





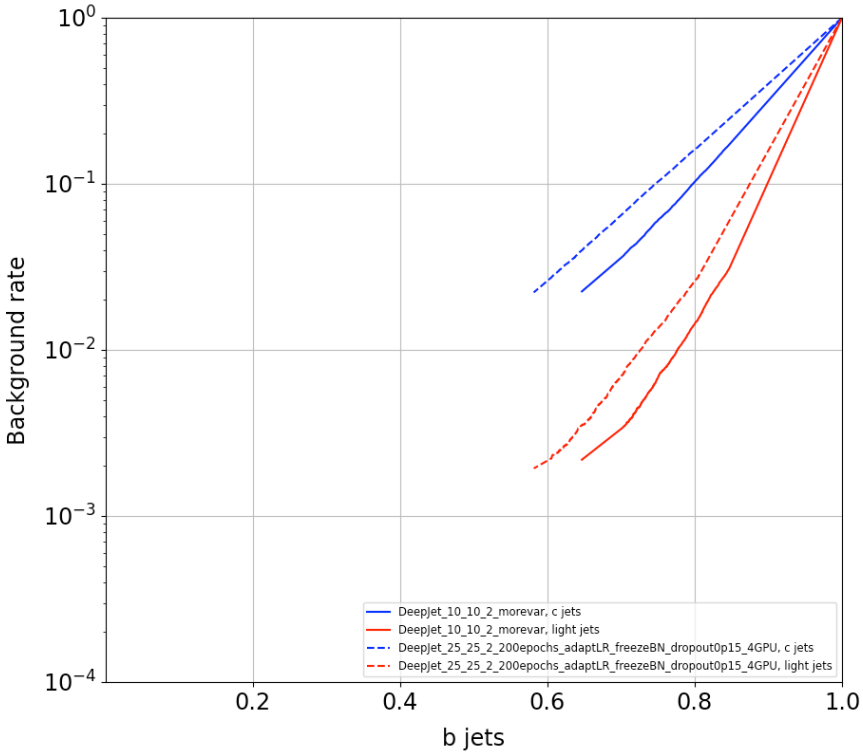
dropout rate 0.1

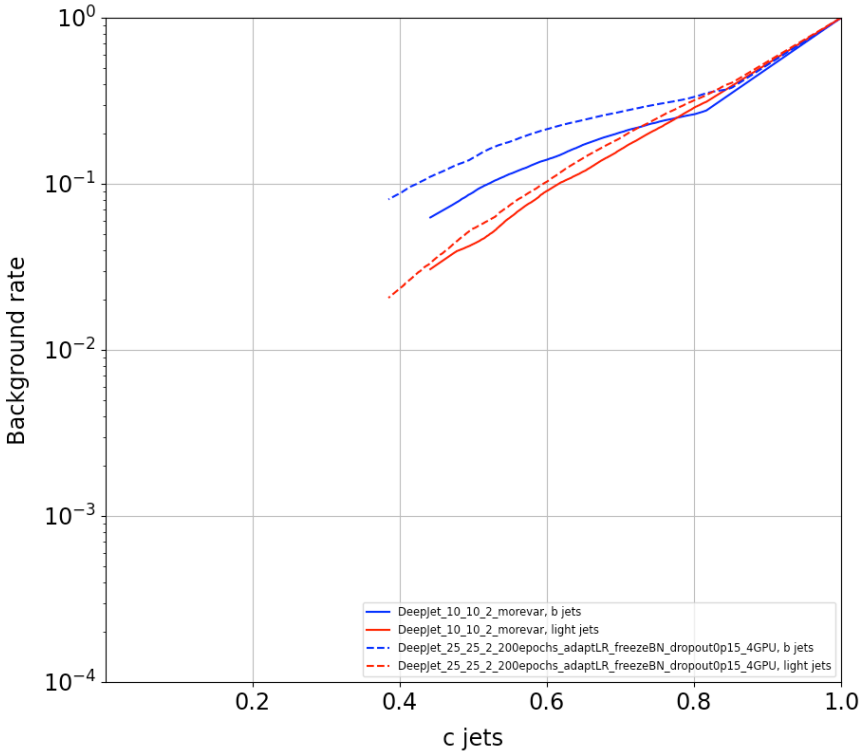




dropout rate 0.15

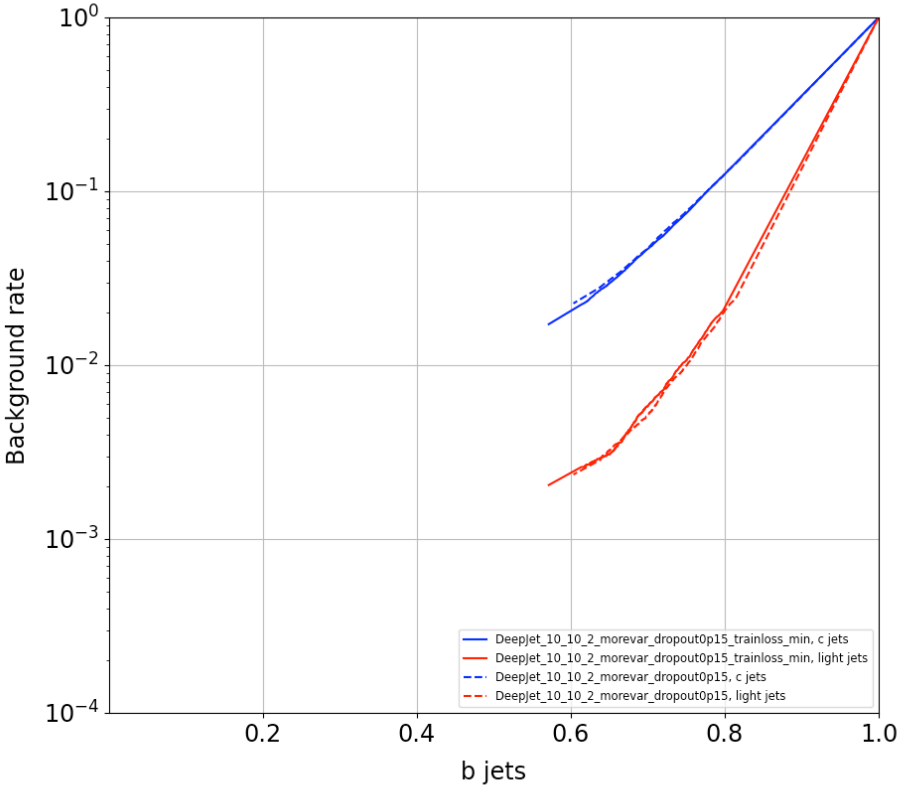
Results: training with more variables vs. more constituents

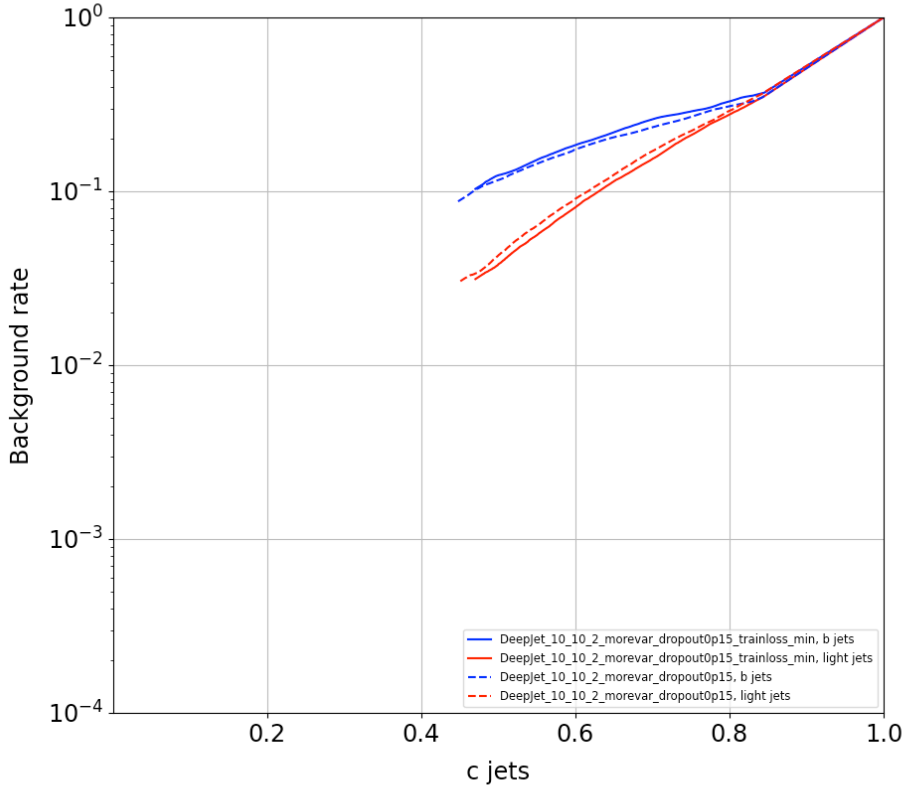




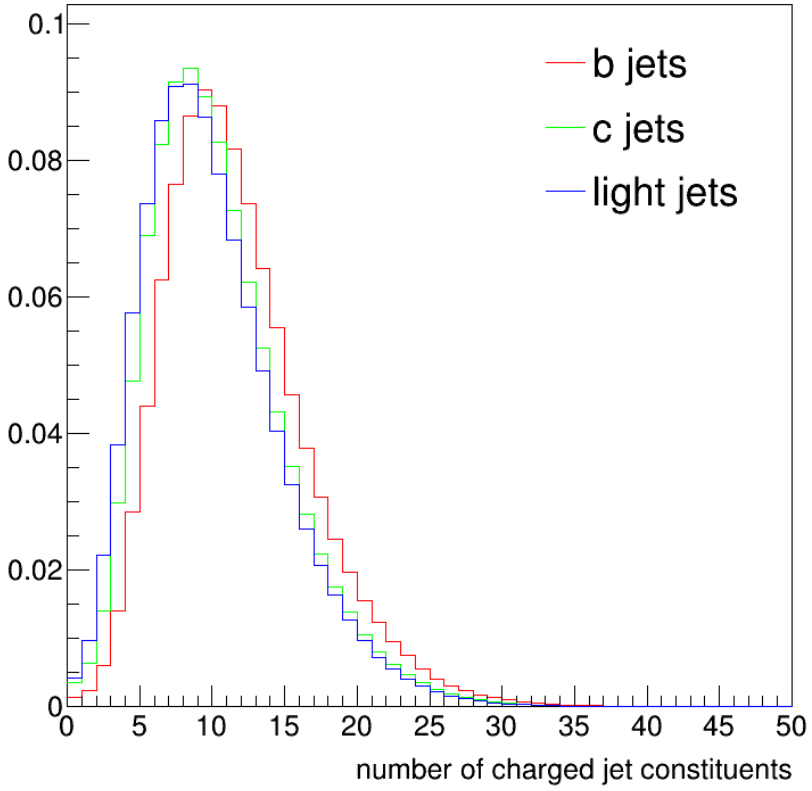
Comparison min training loss vs. max validation accuracy

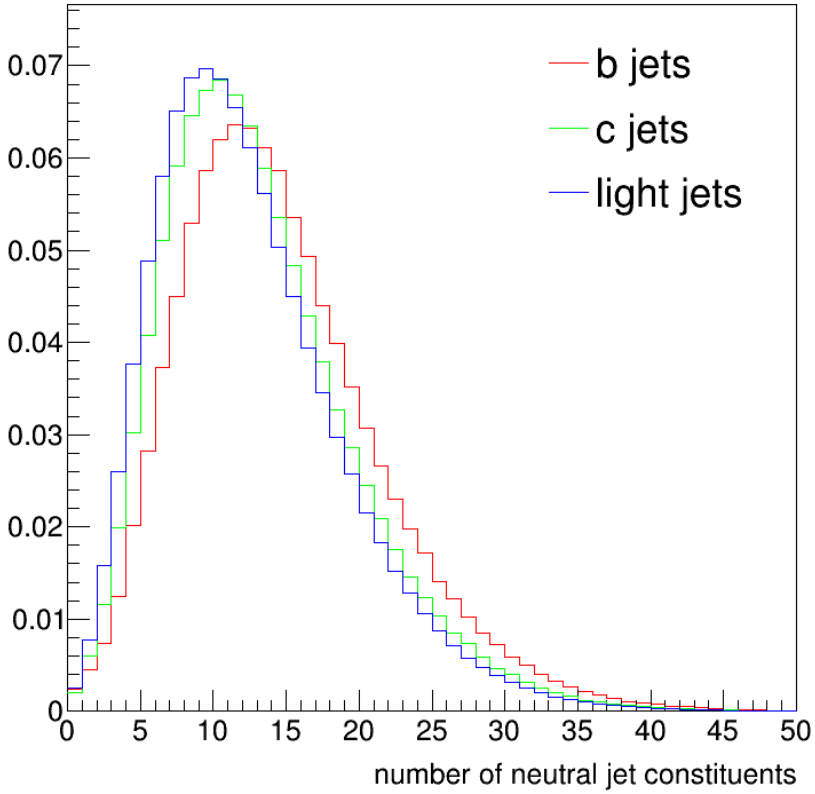
- min training loss: epoch 297
- max validation accuracy: epoch 294



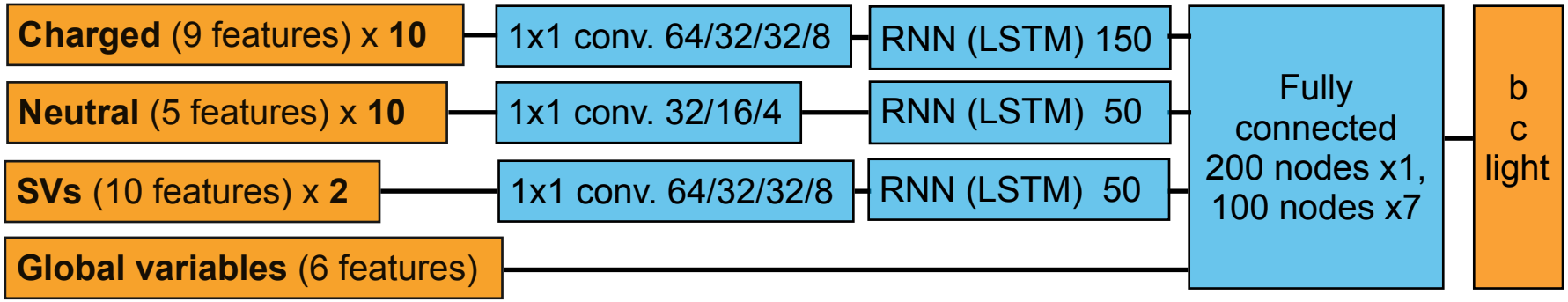


Backup


b jets
c jets
light jets
0.1
0.08
0.06
0.04
0.02
0
0 5 10 15 20 25 30 35 40 45 50
number of charged jet constituents

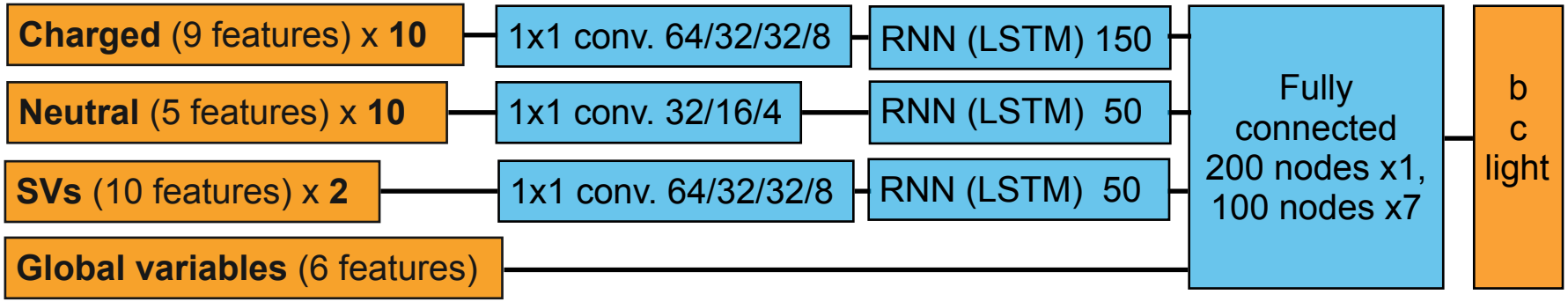

b jets
c jets
light jets
0.07
0.06
0.05
0.04
0.03
0.02
0.01
0
0 5 10 15 20 25 30 35 40 45 50
number of neutral jet constituents

DeepJet: architecture & training



- ordering of particle objects by
  - impact parameter significance for charged jet constituents
  - shortest angular distance to a secondary vertex (by momentum if there is no secondary vertex) for neutral jet constituents
  - flight distance significance for secondary vertices

DeepJet: architecture & training



- batch normalization, dropout (0.1)
- activation functions: relu / softmax (last layer)
- cross entropy loss
- optimizer: Adam
- batch size: 50
- learning rate: 0.0003
- 75% training, 12.5% validation, 12.5% test
- number of epochs: 100
- balanced input: select c/light jets randomly to get same number of b,c,light jets
- Xavier weight initialization

Input features

jets:

- jet rapidity, jet momentum, number of charged jet constituents, number of neutral jet constituents, number of secondary vertices, number of pseudo-vertices

charged jet constituents:

- momentum,Θ(track,jet), ΔR(track, SV), d0 significance, Z0 significance, 3D IP significance, track reconstructed in PV, lepton momentum relative to jet, kaon-ness of charged particles

neutral jet constituents:

- ΔR(neutral,SV), ΔR(neutral,jet), neutral HCAL fraction,momentum/jet momentum, is photon?

secondary vertices:

- secondary vertex mass, number of tracks in SV, χ2/ndf, ΔR(SV,closest jet), SV energy/ jet energy, cosine of the angle between the secondary vertex flight direction and the direction of the secodary vertex momentum, d0 significance, Z0 significance, 3D IP, 3D IP significance

Comparison different dropout rates