Machine Learning Techniques for Heavy Flavour Identification

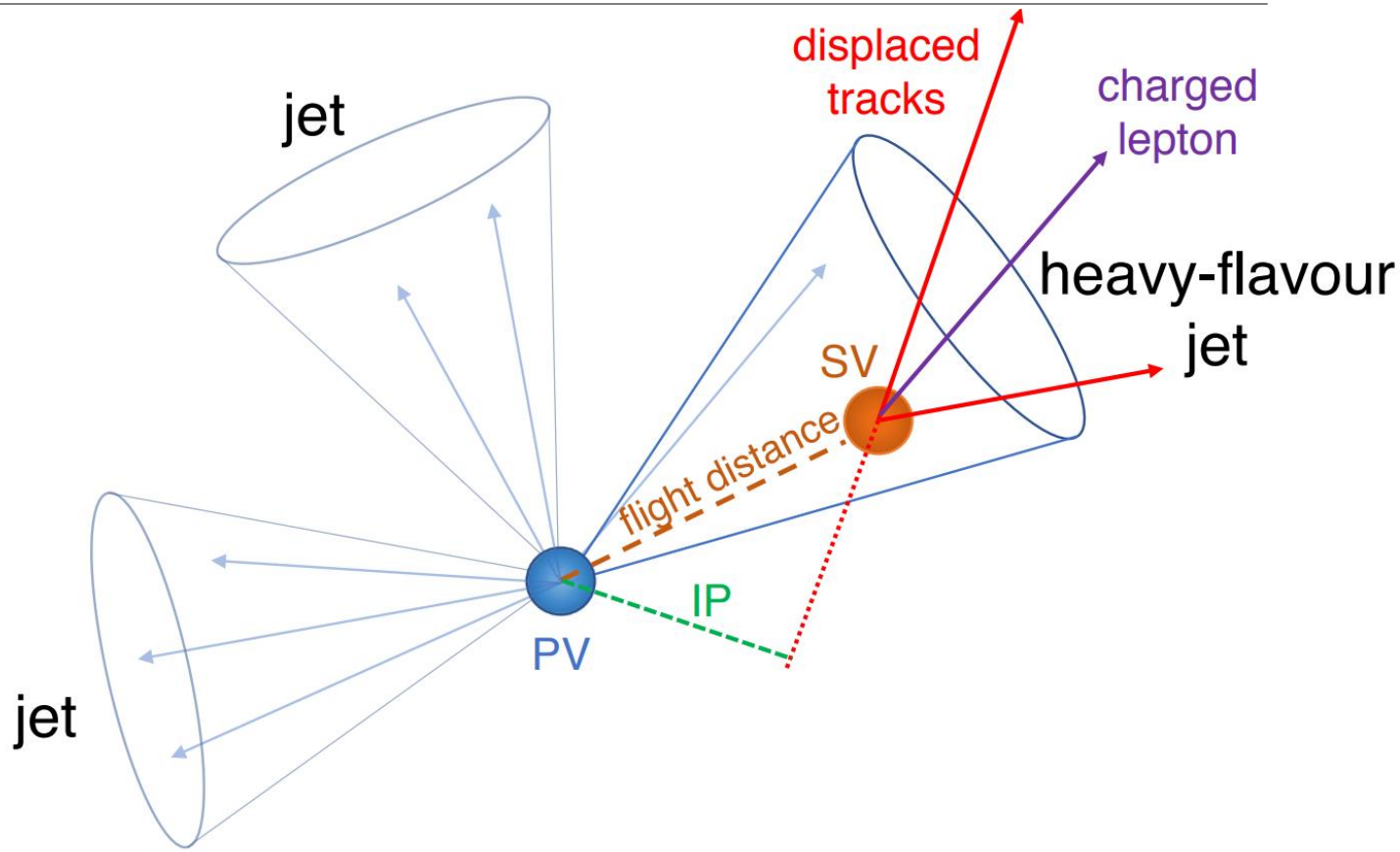
EMIL BOLS

ON BEHALF OF THE CMS COLLABORATION

b-tagging

JINST 13 (2018) P05011

- Several signatures of b-jets
- Displaced tracks
- Secondary vertices
- Soft leptons



b-tagging

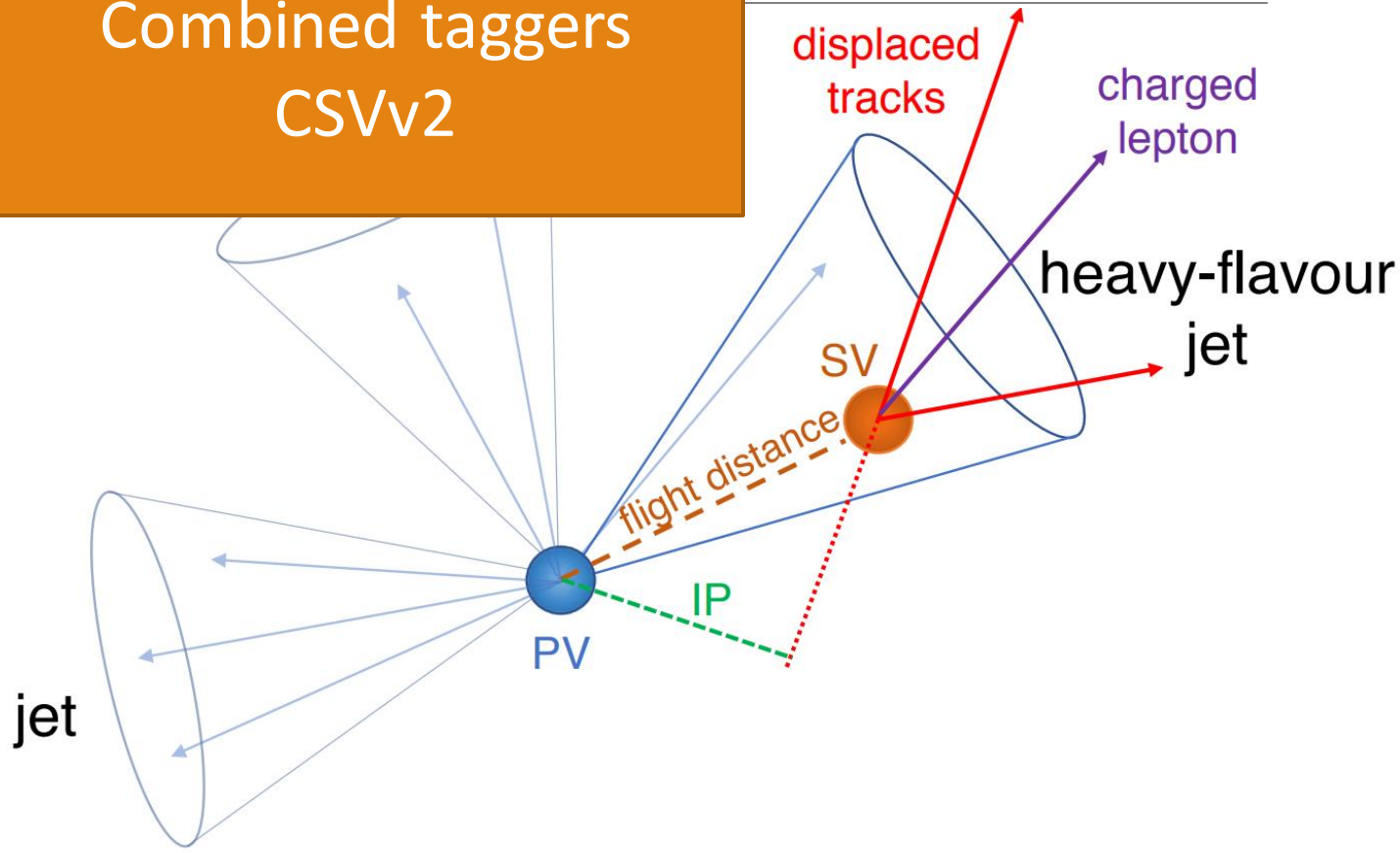
- Several signatures of b-jets

- Displaced tracks
- Secondary vertices

- Soft leptons

Combined taggers
CSVv2

JINST 13 (2018) P05011



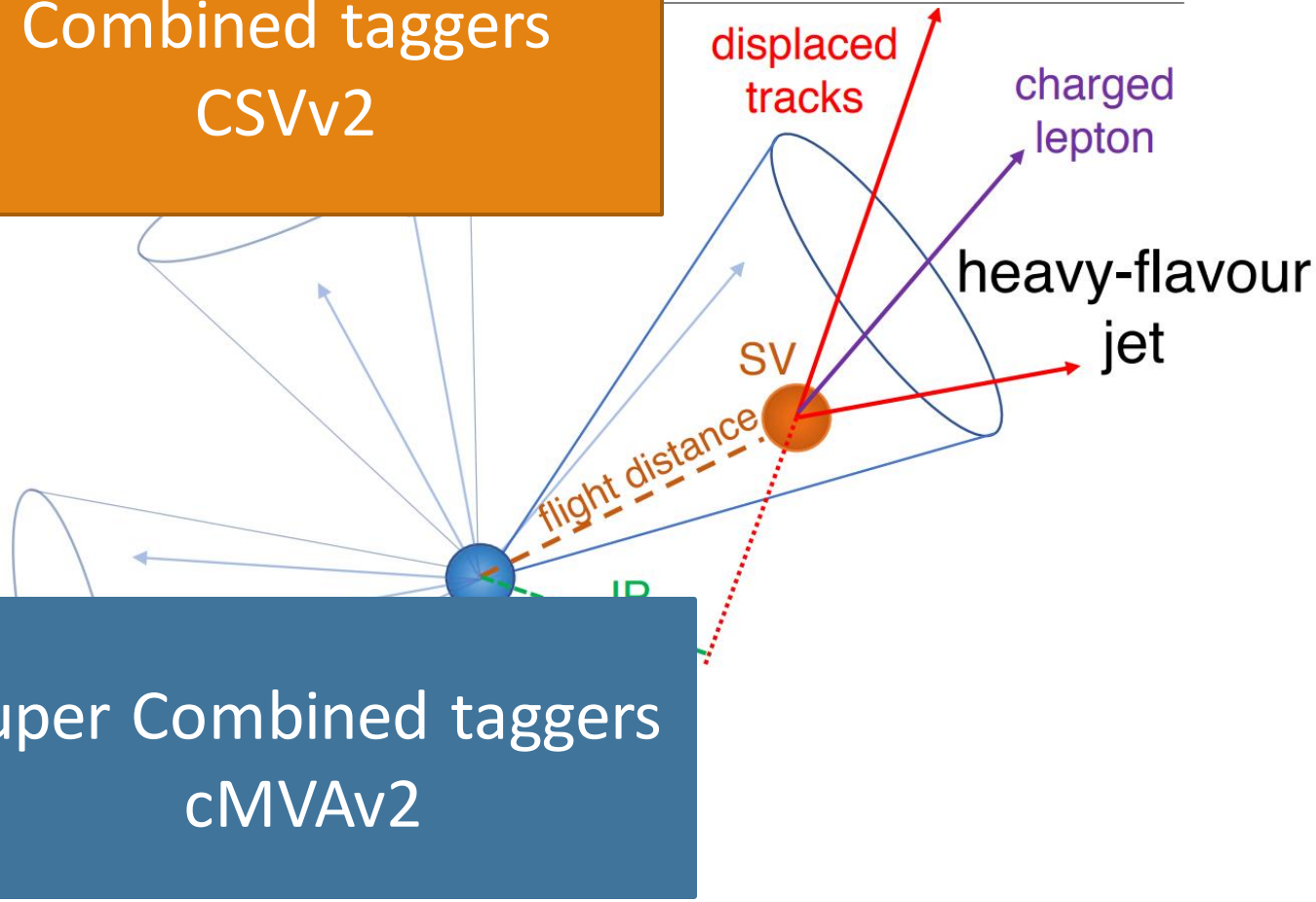
b-tagging

JINST 13 (2018) P05011

- Several signatures of b-jets
  - Displaced tracks
  - Secondary vertices
  - Soft leptons

Combined taggers
CSVv2

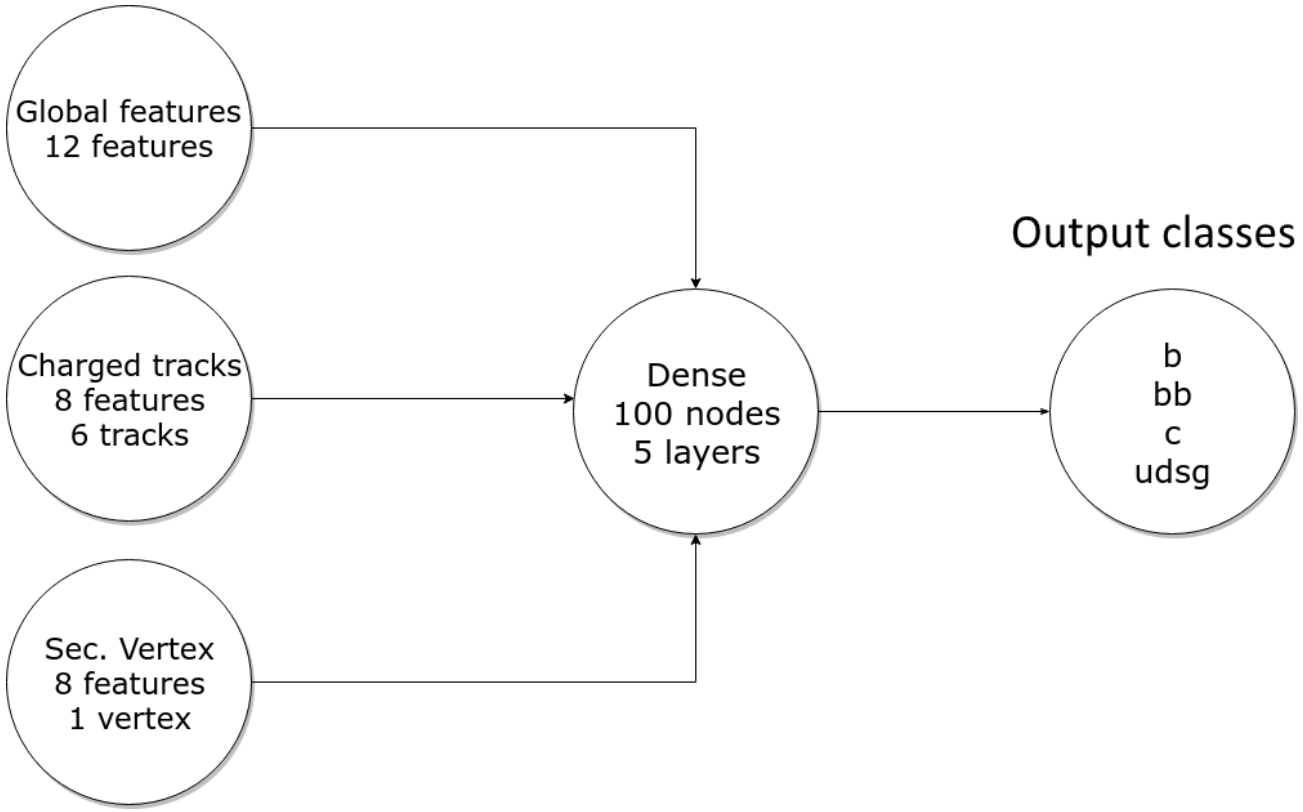
Super Combined taggers
cMVAv2

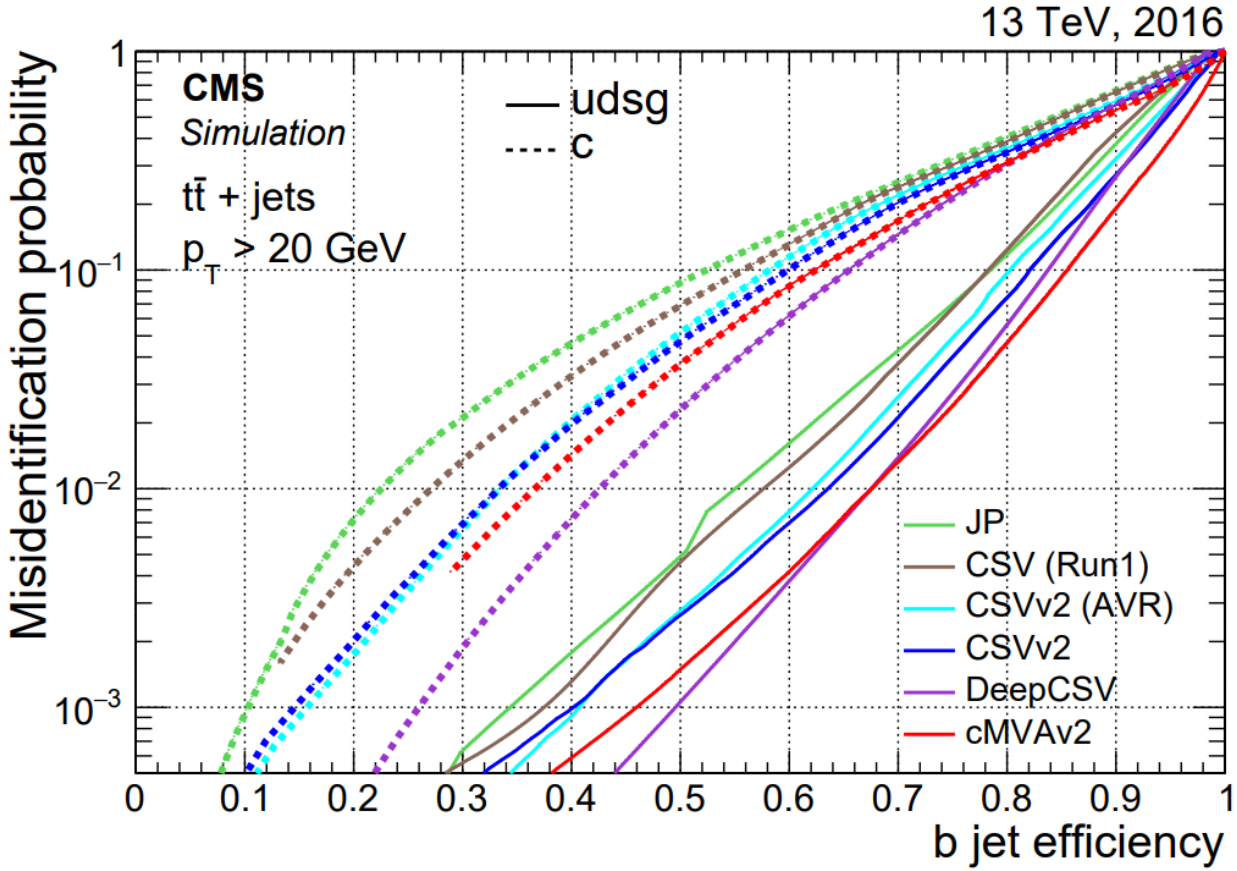


DeepCSV

JINST 13 (2018) P05011

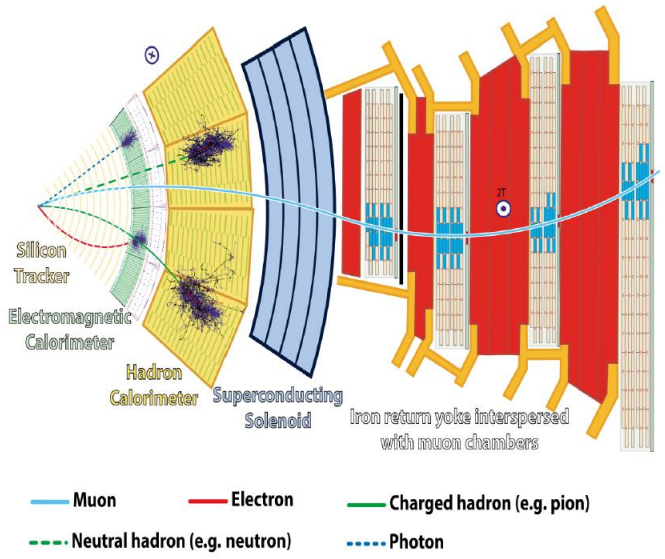
- DNN with similar input as CSV

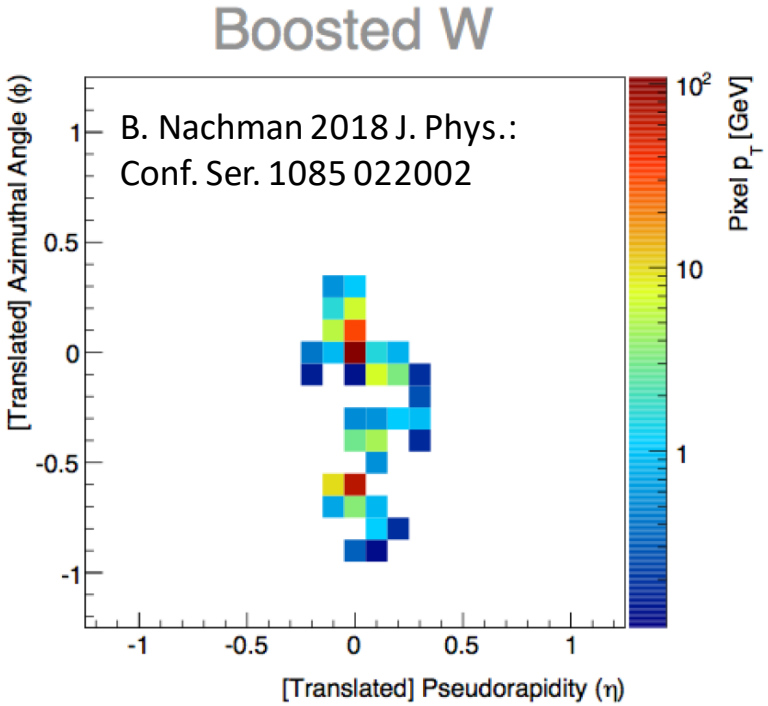




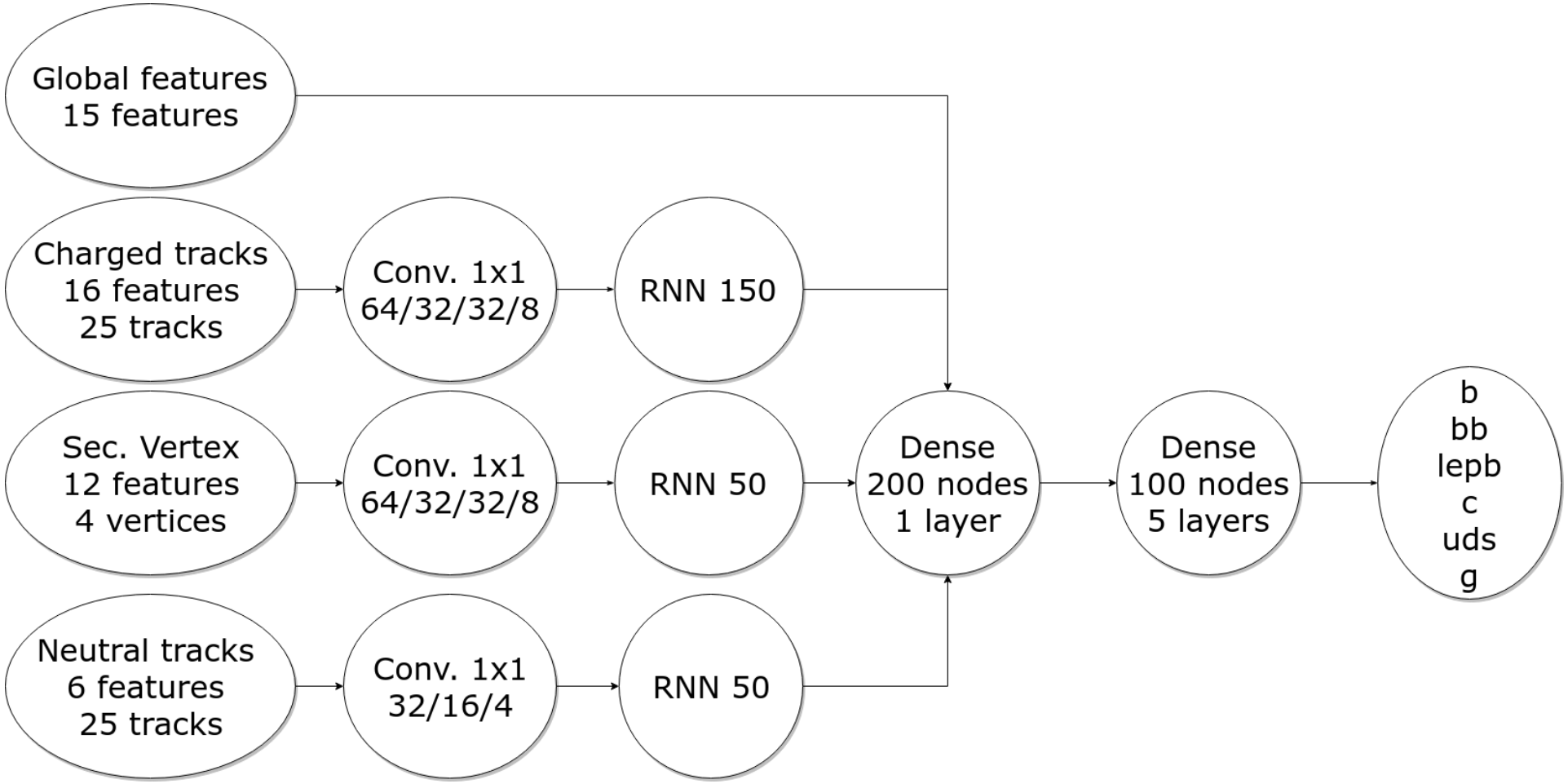
Trying more complex architectures

- Success of convolutional neural networks for computer vision
- Should jets be represented as images?
- Challenging for heavy flavour ID - CMS tracking information is complex
- Particle based architecture instead

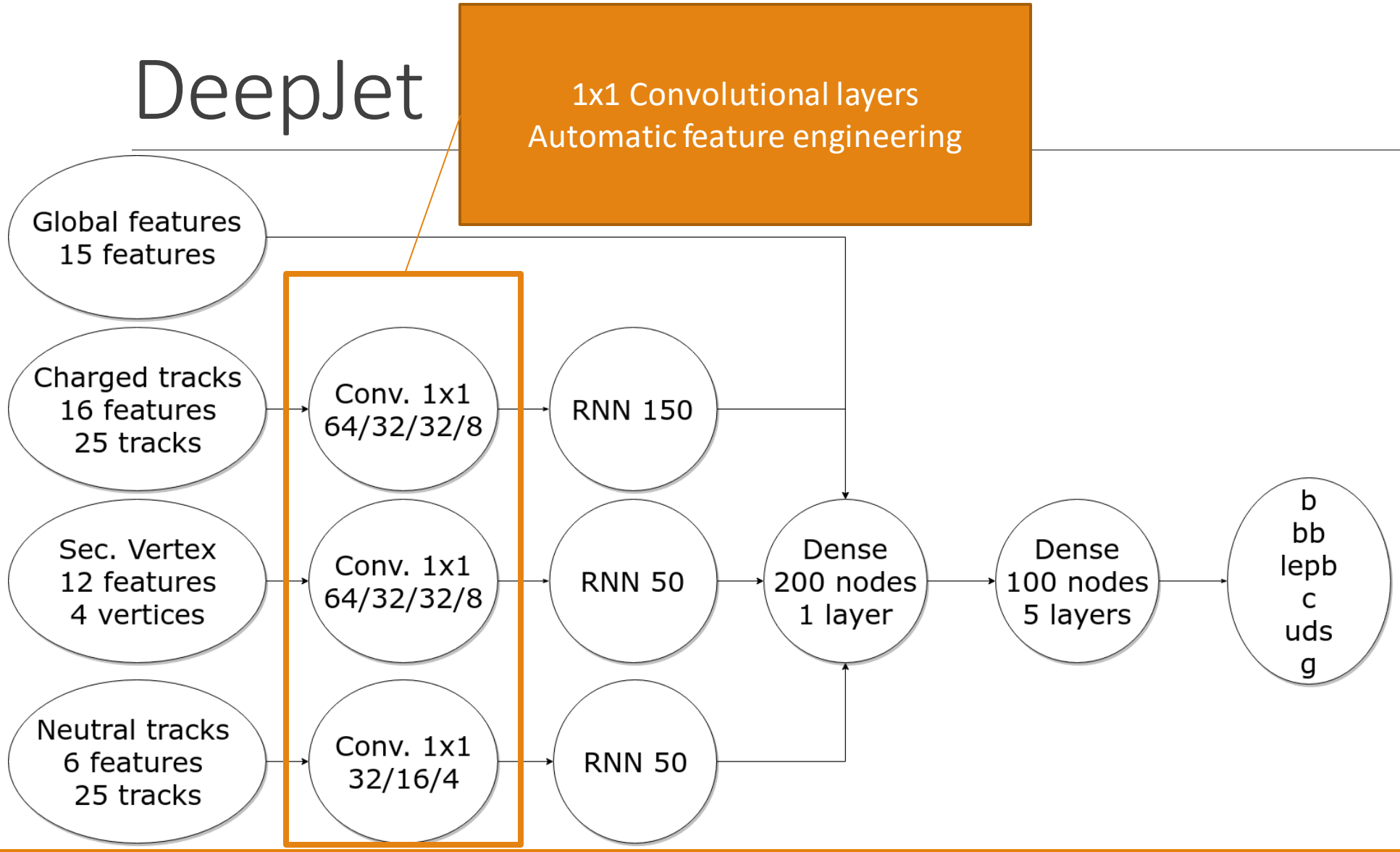




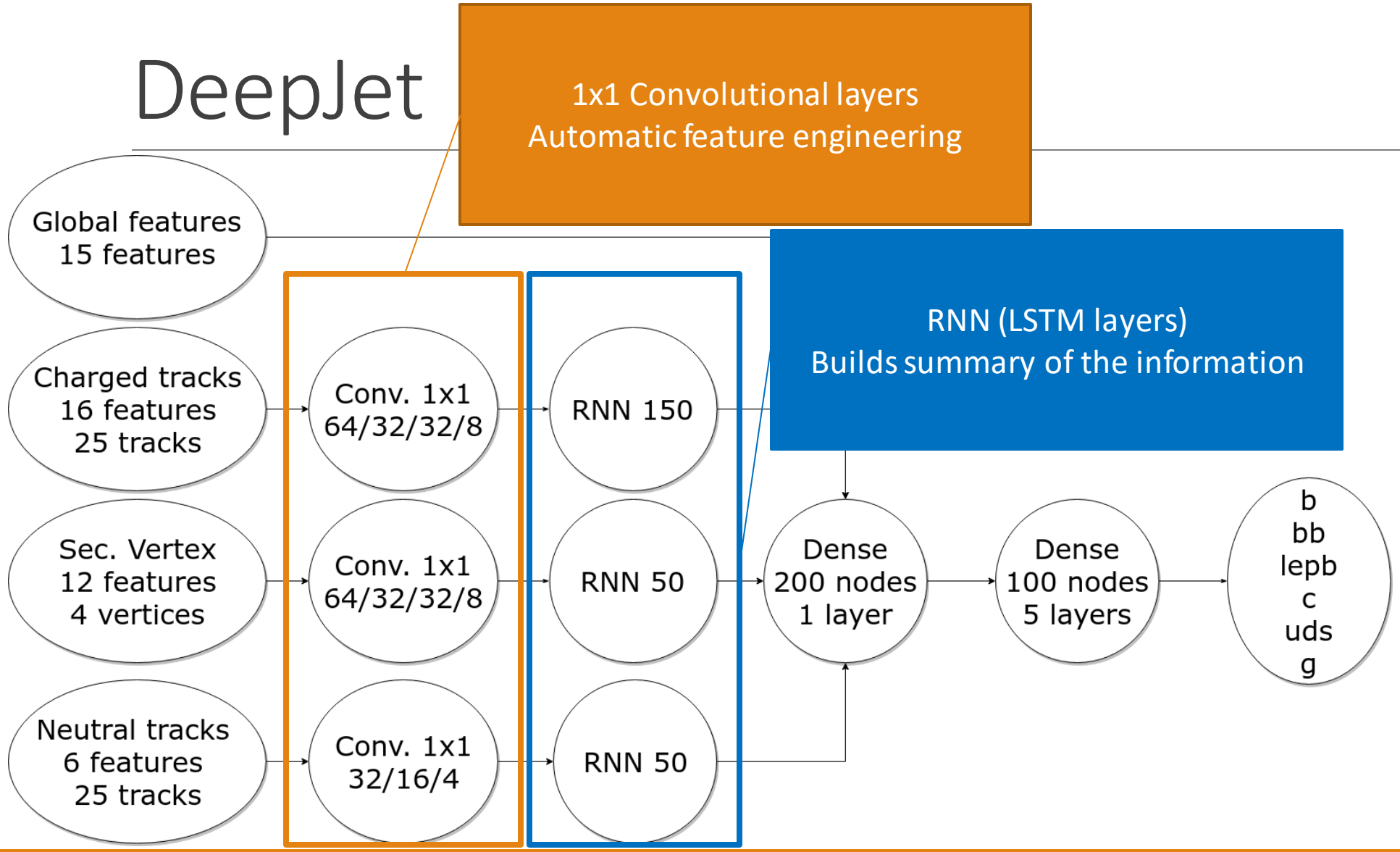
DeepJet


Global features
15 features
Charged tracks
16 features
25 tracks
Conv. 1x1
64/32/32/8
RNN 150
Sec. Vertex
12 features
4 vertices
Conv. 1x1
64/32/32/8
RNN 50
Neutral tracks
6 features
25 tracks
Conv. 1x1
32/16/4
RNN 50
Dense
200 nodes
1 layer
Dense
100 nodes
5 layers
b
bb
lepb
c
uds
g

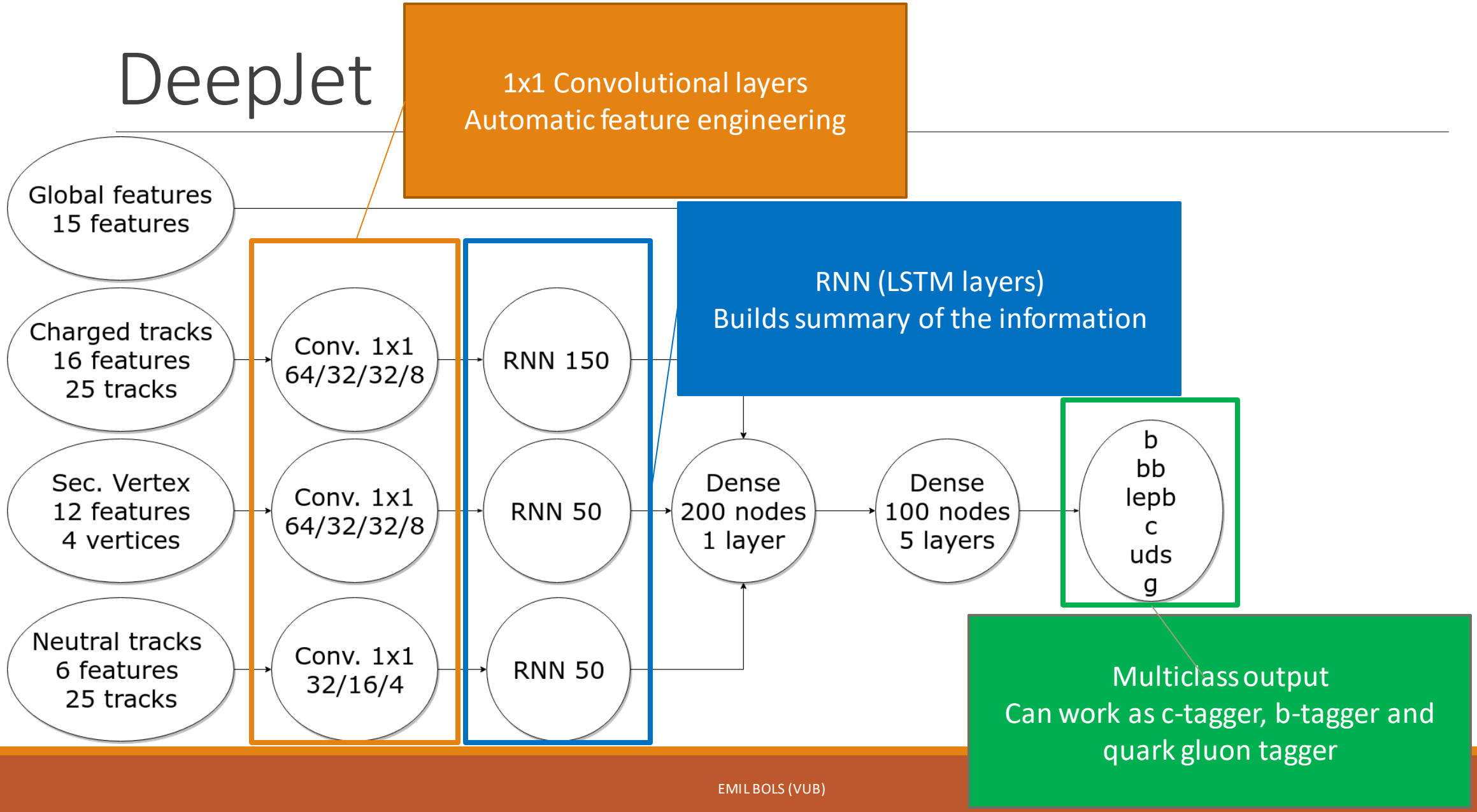
DeepJet


1x1 Convolutional layers
Automatic feature engineering
Global features
15 features
Charged tracks
16 features
25 tracks
Sec. Vertex
12 features
4 vertices
Neutral tracks
6 features
25 tracks
Conv. 1x1
64/32/32/8
Conv. 1x1
64/32/32/8
Conv. 1x1
32/16/4
RNN 150
RNN 50
RNN 50
Dense
200 nodes
1 layer
Dense
100 nodes
5 layers
b
bb
lepb
c
uds
g

DeepJet


1x1 Convolutional layers
Automatic feature engineering
RNN (LSTM layers)
Builds summary of the information
Global features
15 features
Charged tracks
16 features
25 tracks
Sec. Vertex
12 features
4 vertices
Neutral tracks
6 features
25 tracks
Conv. 1x1
64/32/32/8
Conv. 1x1
64/32/32/8
Conv. 1x1
32/16/4
RNN 150
RNN 50
RNN 50
Dense
200 nodes
1 layer
Dense
100 nodes
5 layers
b
bb
lepb
c
uds
g

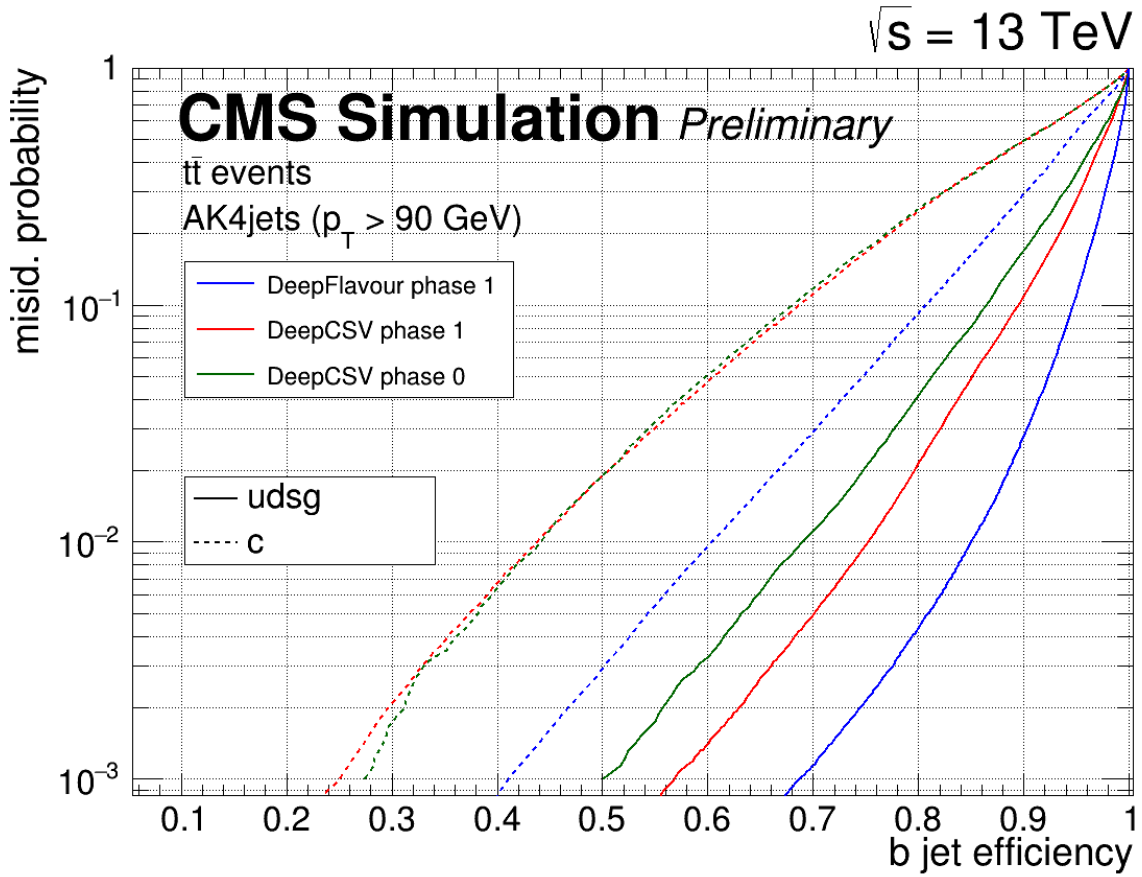
DeepJet


1x1 Convolutional layers
Automatic feature engineering
Global features
15 features
Charged tracks
16 features
25 tracks
Sec. Vertex
12 features
4 vertices
Neutral tracks
6 features
25 tracks
Conv. 1x1
64/32/32/8
Conv. 1x1
64/32/32/8
Conv. 1x1
32/16/4
RNN 150
RNN 50
RNN 50
RNN (LSTM layers)
Builds summary of the information
Dense
200 nodes
1 layer
Dense
100 nodes
5 layers
b
bb
lepb
c
uds
g
Multiclass output
Can work as c-tagger, b-tagger and
quark gluon tagger

DeepJet

CMS DP-2018/033

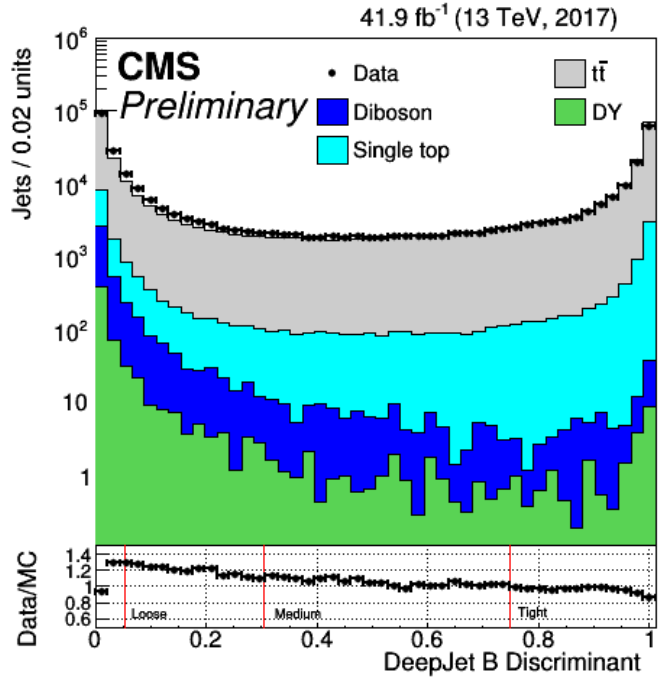
- Large improvement in performance in simulation
- However simulation is never perfect
- Will this gain translate into data?

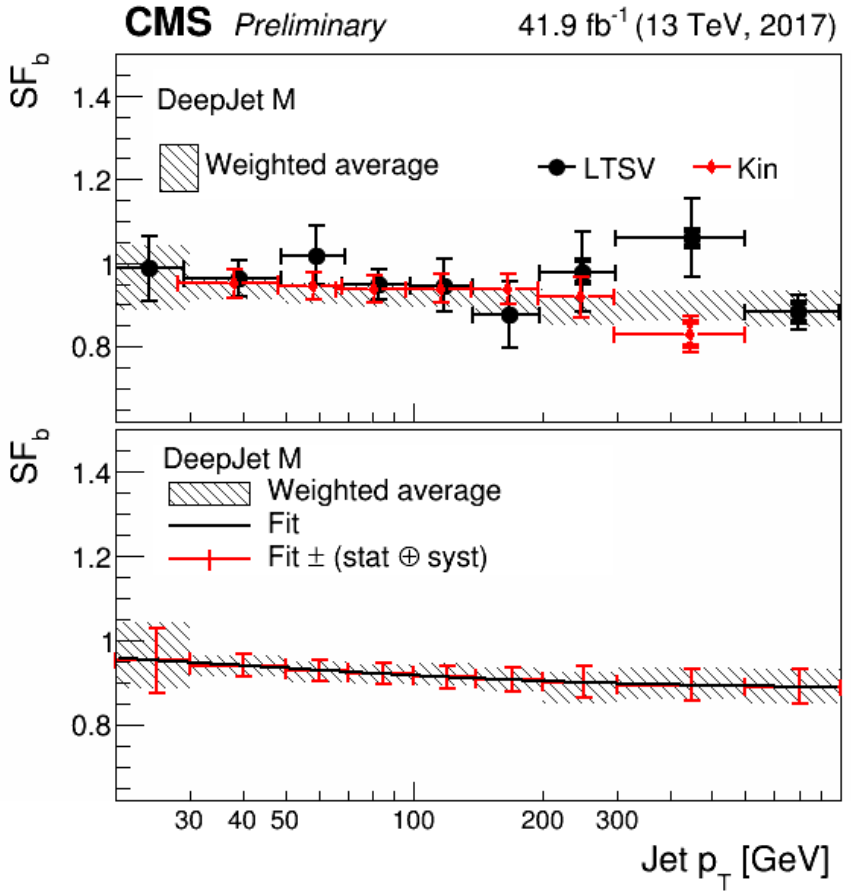


DeepJet

CMS DP-2018/058

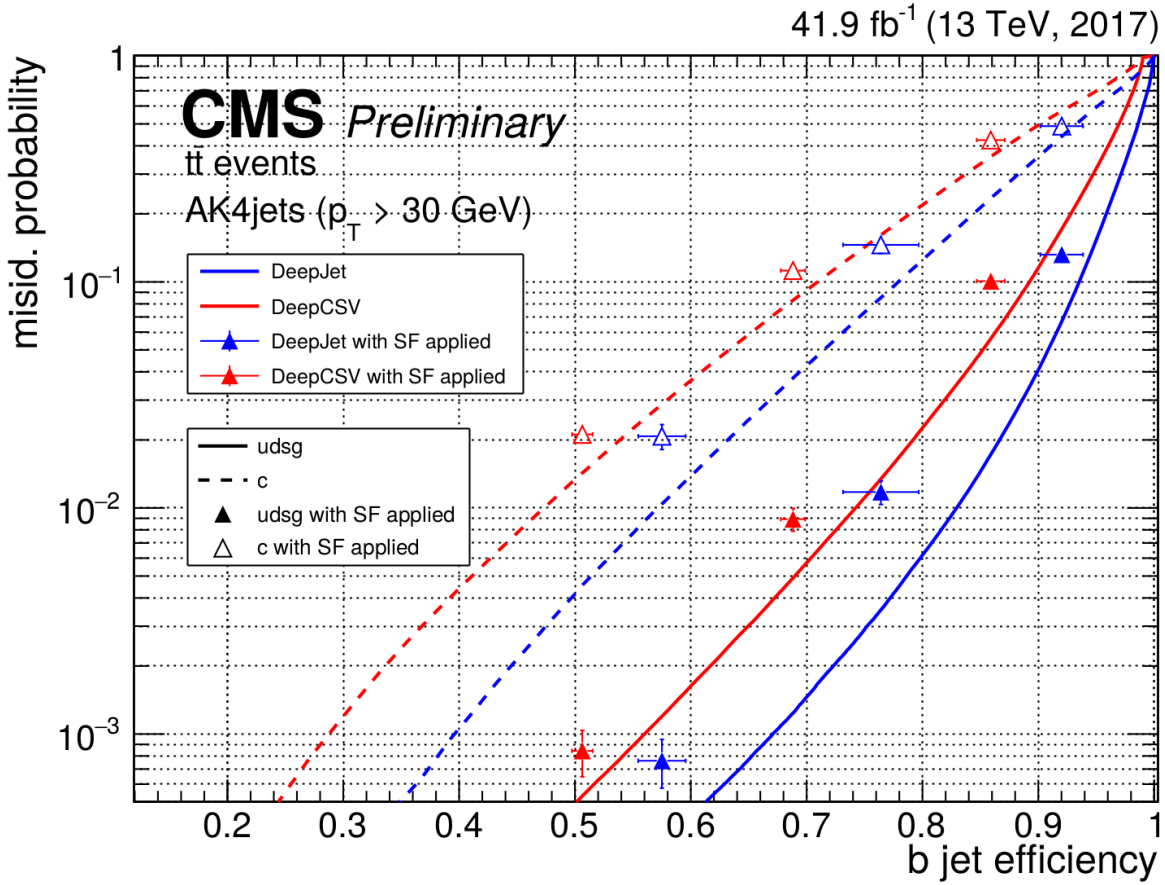
- Simulation to data scale factors have been derived for DeepJet for the first time

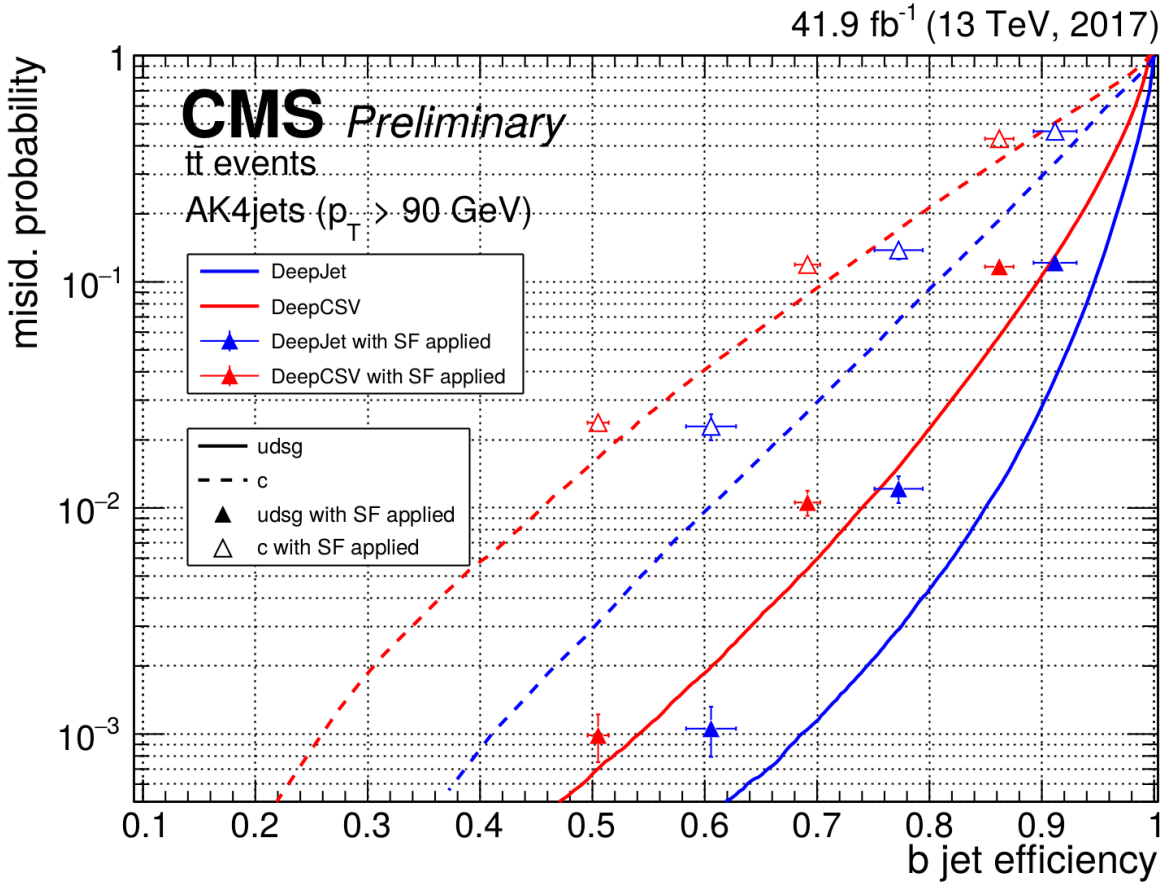




DeepJet

CMS DP-2018/058

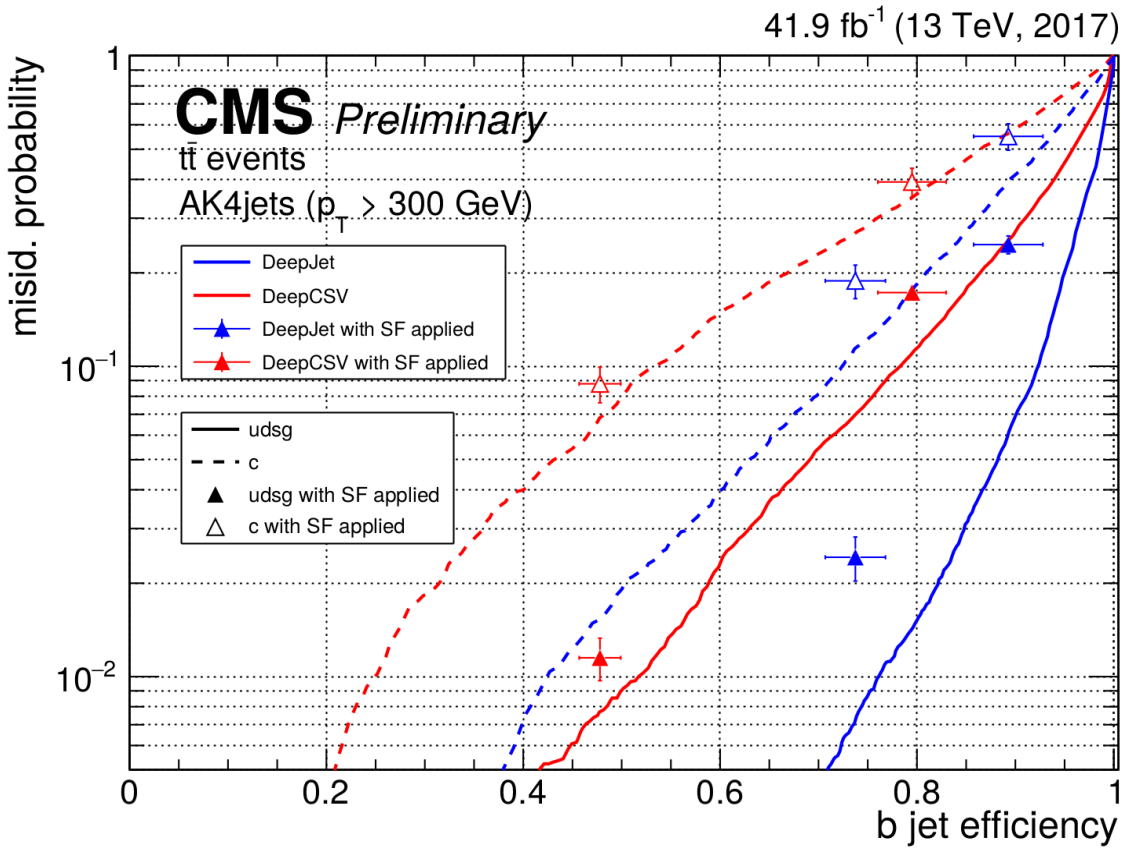




DeepJet

CMS DP-2018/058

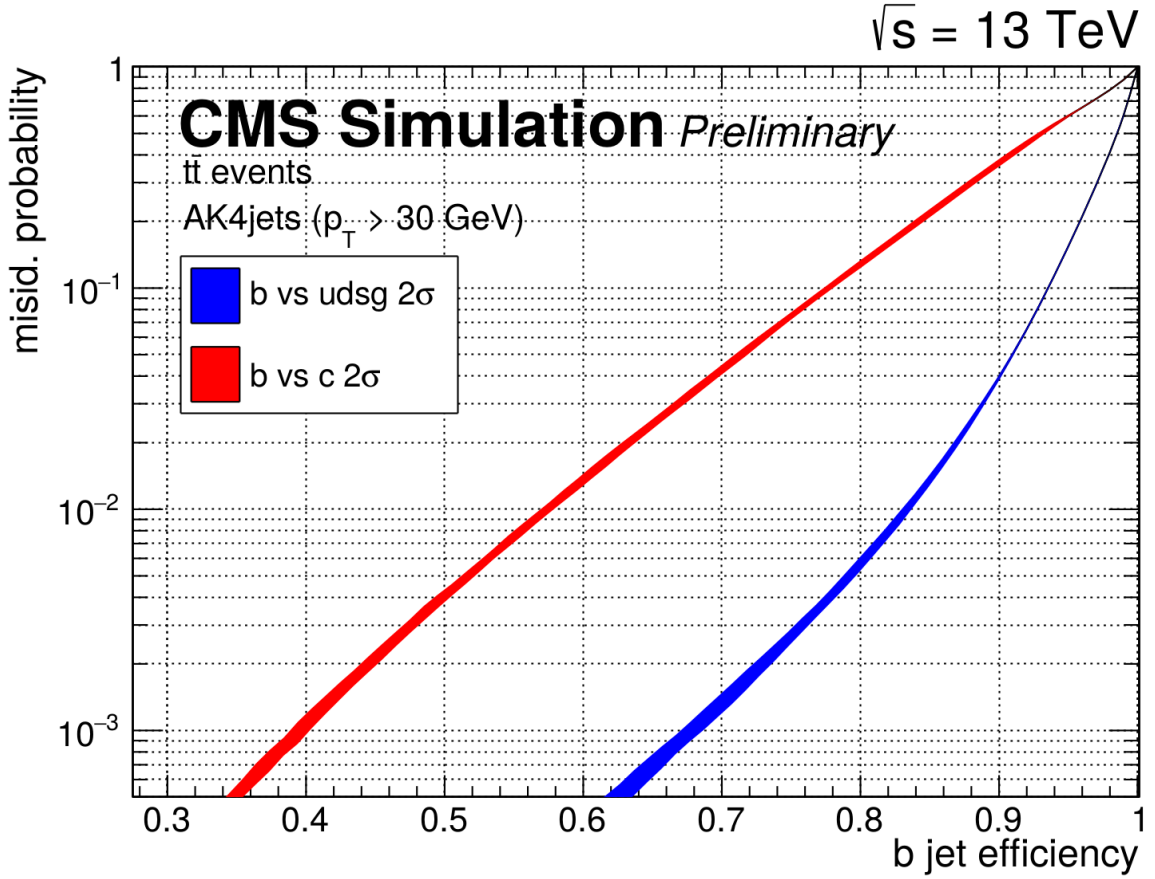
- b-tagging in the very boosted regime with DeepJet
- Since DeepJet has minimal track selection, it can fully utilize the jet information
- A large gain in performance



ML studies

CMS DP-2018/058

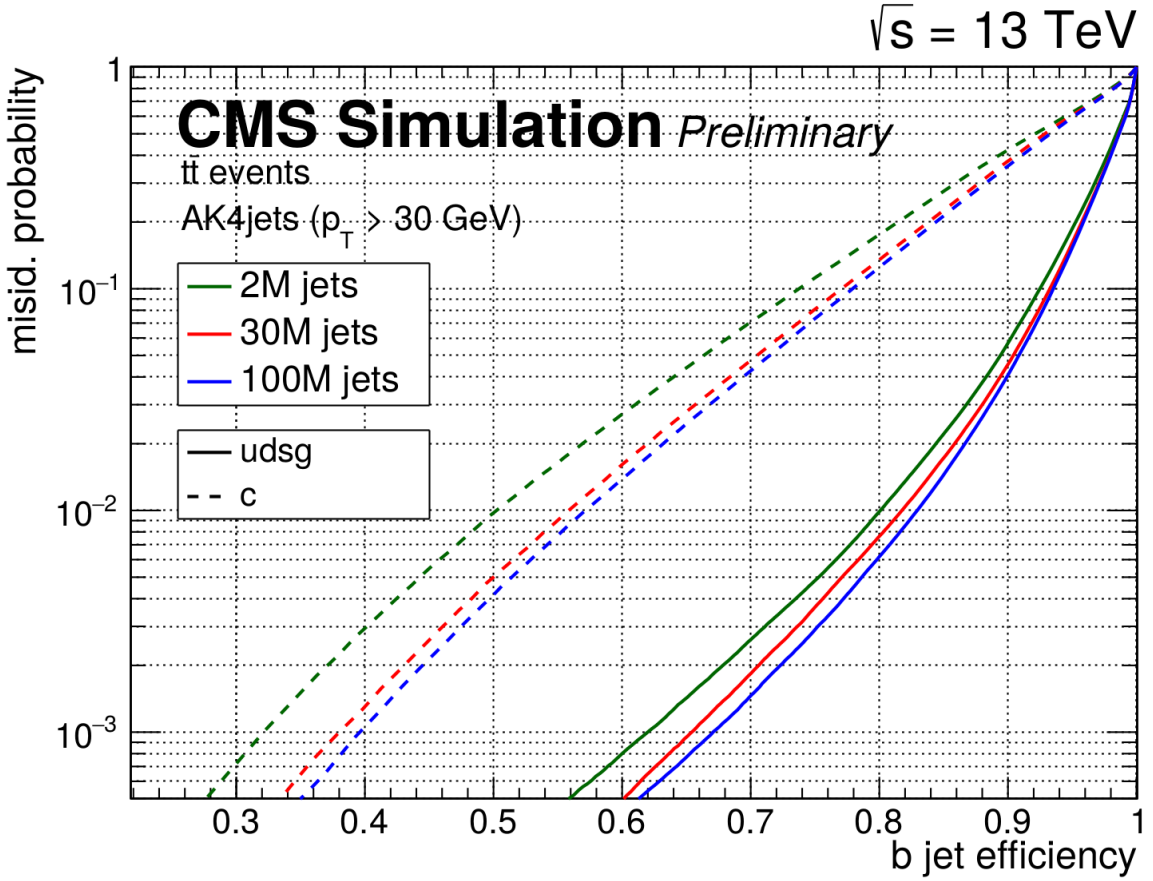
- Several ongoing studies of the properties of the network
- Dependence on random initialization of weights



ML studies

CMS DP-2018/058

- Several ongoing studies of the properties of the network
- Dependence on random initialization of weights
- Dependence on training sample datasize

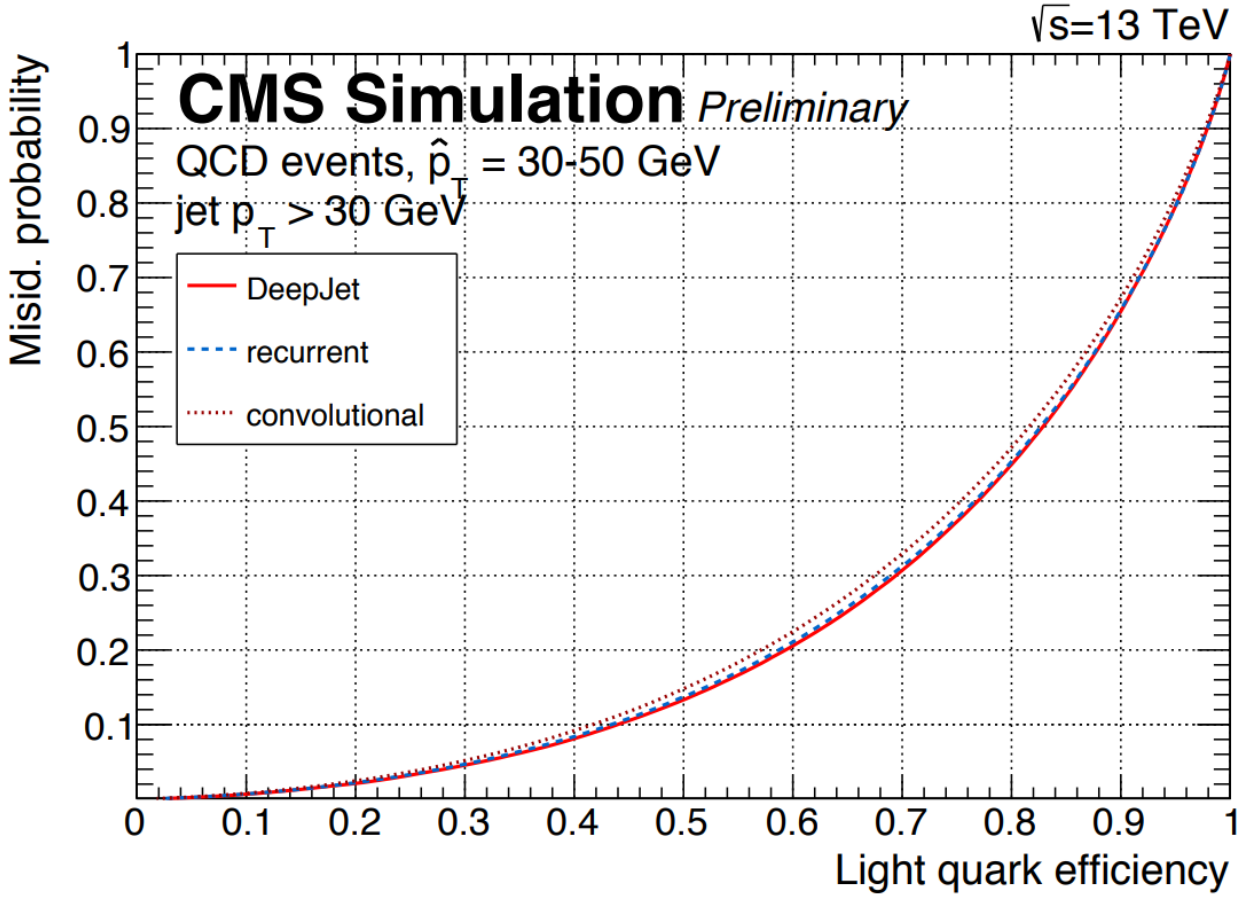


Summary

- The DeepJet algorithm has been commissioned in data for the first time.
- Large gain in performance in data compared to previous taggers, both in the inclusive sample and in the boosted regime.
- Several ongoing studies to optimize the algorithm, such as model distillation, pruning and interpretation.

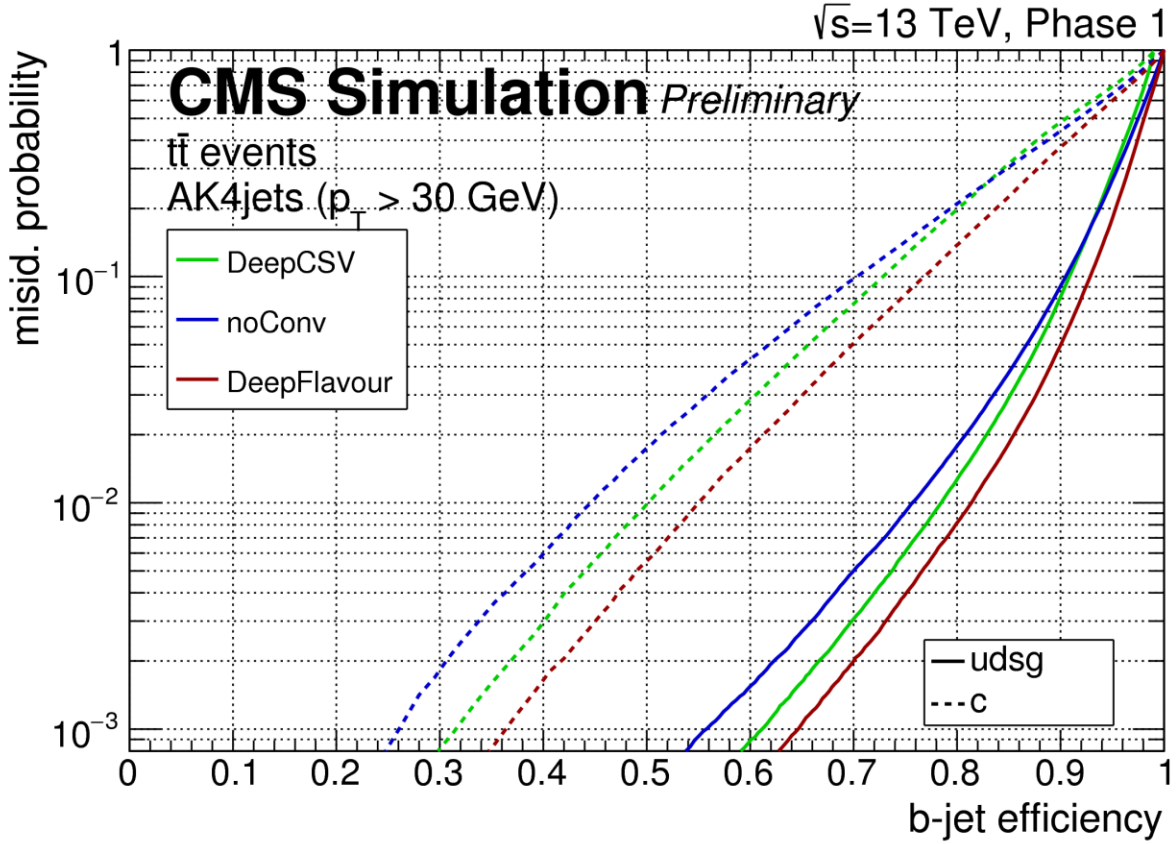
Backup

QG performance



CMS DP-2017/027

Without conv. layer



CMS DP-2017/013