Data

- Training : ~15k b-jets and dark-jets
- Testing : ~15k b-jets and dark-jets
- For all dark jets: rinv = 0 and dark pion decay length = 0.5 mm

Features

- Track information only : [pt, phi, eta, deltaR, D0, DZ]
- Preprocessing: pt->pt/(track-pt), boost and global scaling.

Models

LSTM

Layer (type)	Output Shape	Param #
lstm_4 (LSTM)	(None, 50)	10600
dense_23 (Dense)	(None, 16)	816
dense_24 (Dense)	(None, 1)	17

Total params: 11,433
Trainable params: 11,433
Non-trainable params: 0

CNN

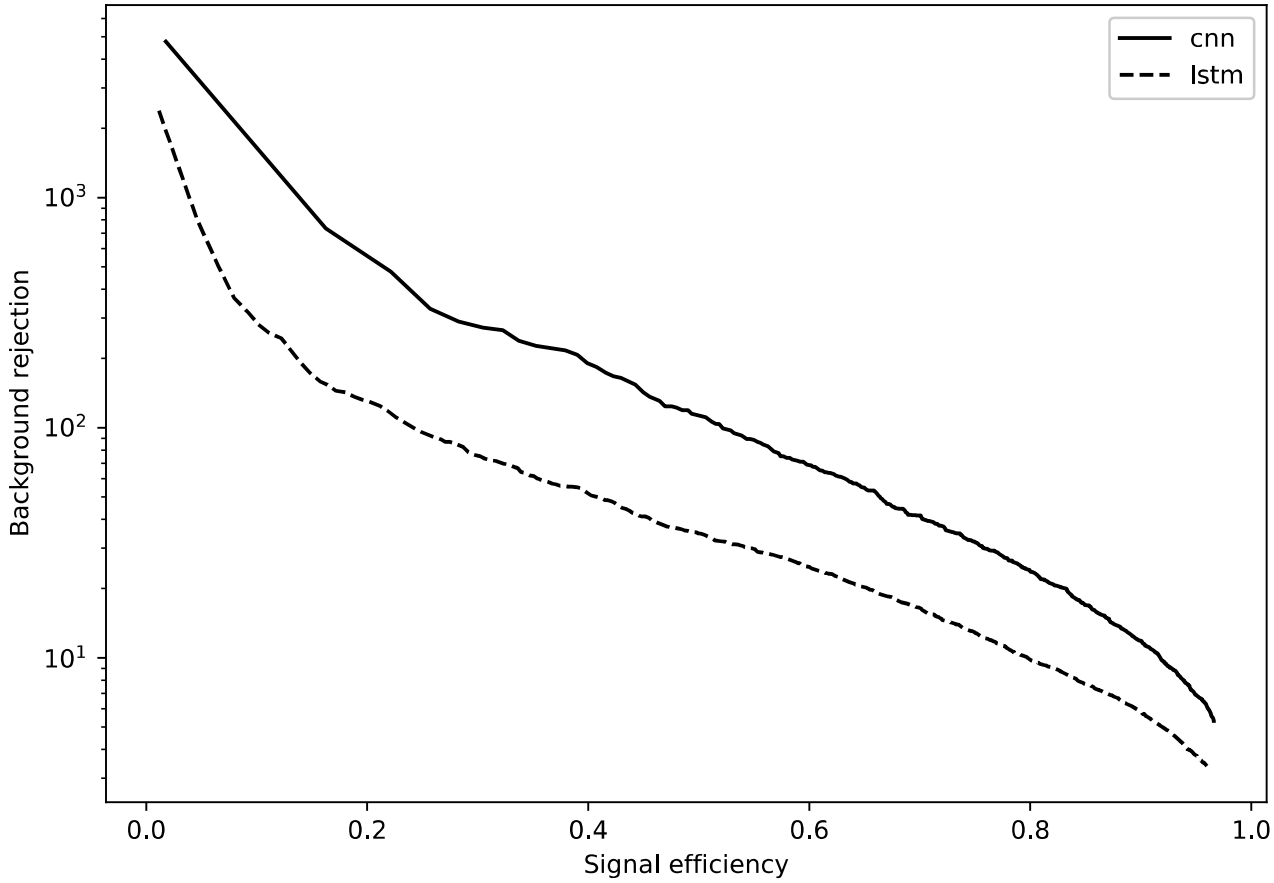
Layer (type)	Output Shape	Param #
conv1d_16 (Conv1D)	(None, 30, 128)	896
conv1d_17 (Conv1D)	(None, 30, 64)	24640
dropout_8 (Dropout)	(None, 30, 64)	0
max_pooling1d_8 (MaxPooling1	(None, 15, 64)	0
conv1d_18 (Conv1D)	(None, 15, 32)	6176
conv1d_19 (Conv1D)	(None, 15, 32)	3104
dropout_9 (Dropout)	(None, 15, 32)	0
max_pooling1d_9 (MaxPooling1	(None, 7, 32)	0
flatten_4 (Flatten)	(None, 224)	0
dense_20 (Dense)	(None, 16)	3600
dense_21 (Dense)	(None, 16)	272
dense_22 (Dense)	(None, 1)	17

Total params: 38,705
Trainable params: 38,705
Non-trainable params: 0

Study

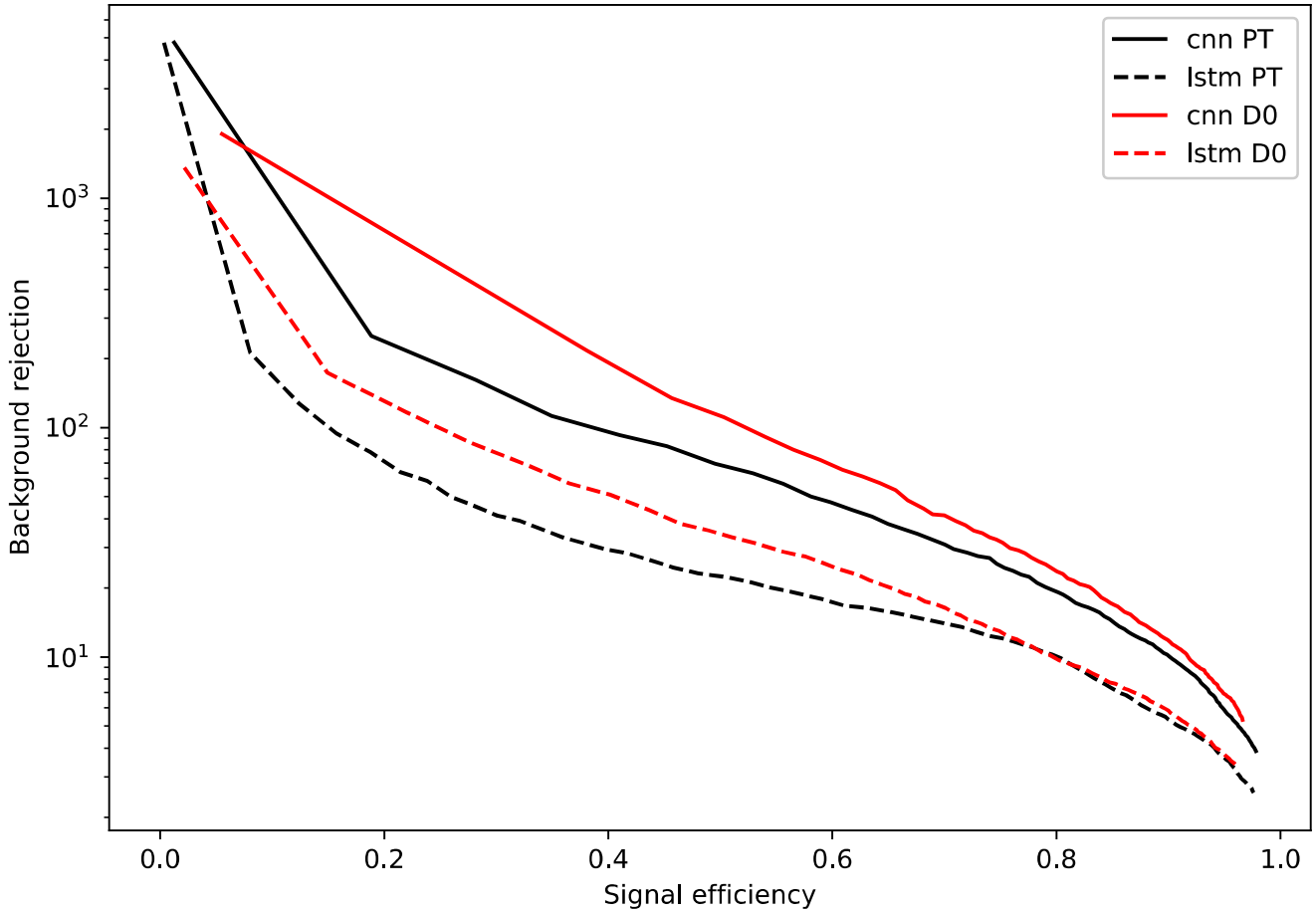
- How many constituents to keep?
- By what feature to sort constituents?
- How to determine the sign of D0?

CNN vs **LSTM**



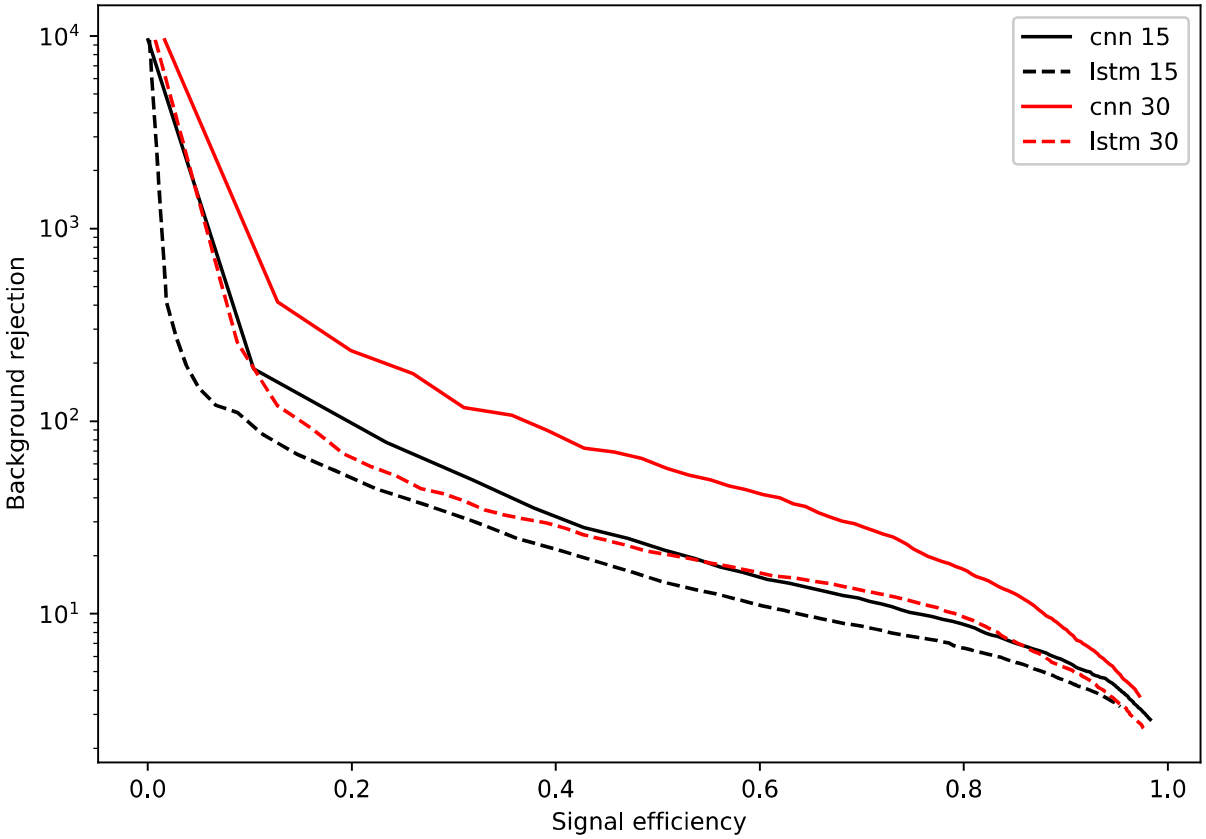
- Sorted by abs(D0).
- 30 Constituents.
- Sign of D0 determined by jet PT.

abs(D0) sort vs **PT sort**



- Slightly better to sort by abs(D0).

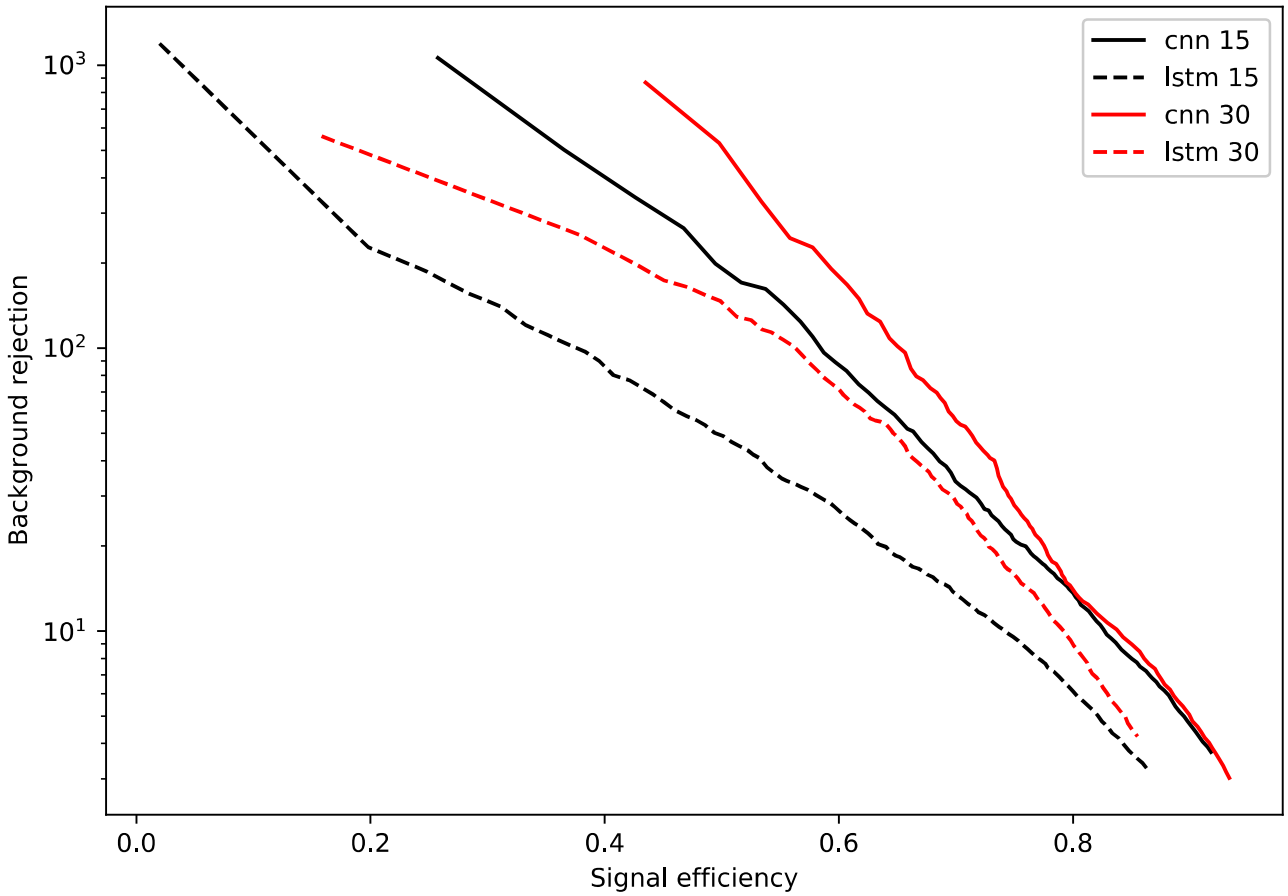
30 constituents vs **15 constituents**
(sort D0)



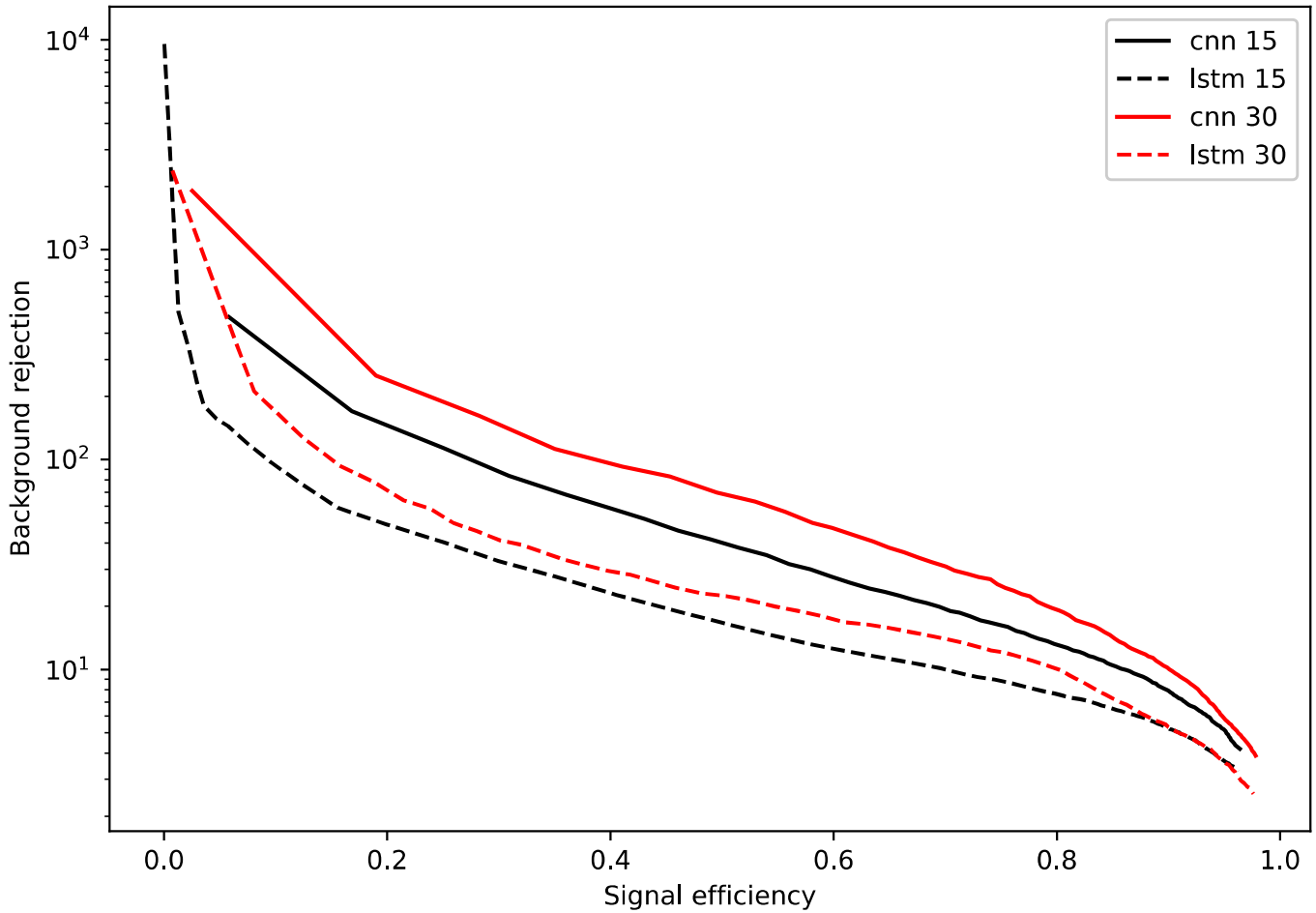
- 30 constituents is better than 15.
- I think this is partially due to multiplicity information.
- This affects CNN more.

30 constituents vs **15 constituents**

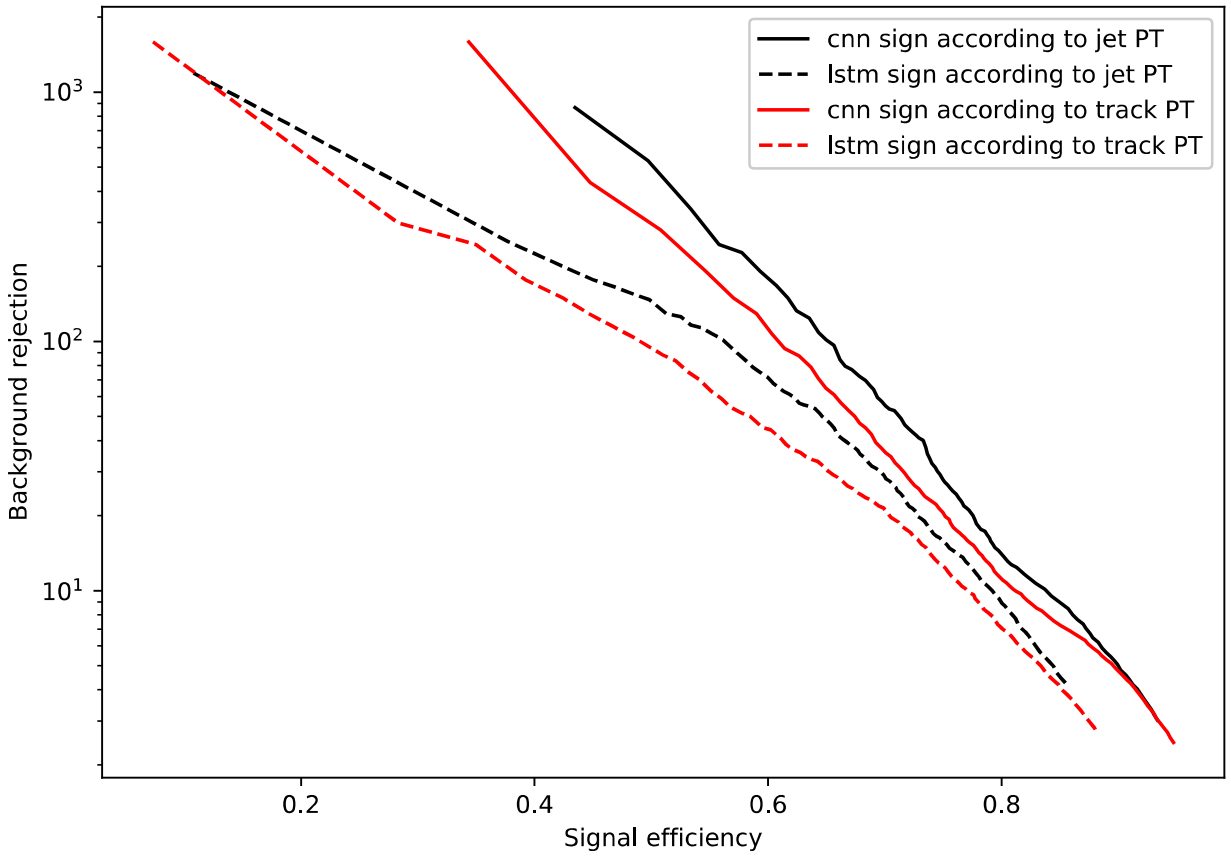
IP features only



30 constituents vs **15 constituents**
(sort PT)



D0 sign wrt jet-PT vs **D0 sign wrt track-PT**
(30 constituents)



- Taking the sign with respect to jet-PT is slightly better.

D0 sign wrt jet-PT vs **D0 sign wrt track-PT**
(15 constituents)