

Track classification in hadronic tau decays with recurrent neural networks at the ATLAS detector

Max Märker

Supervised by:

Dirk Duschinger, Wolfgang Mader, Arno Straessner

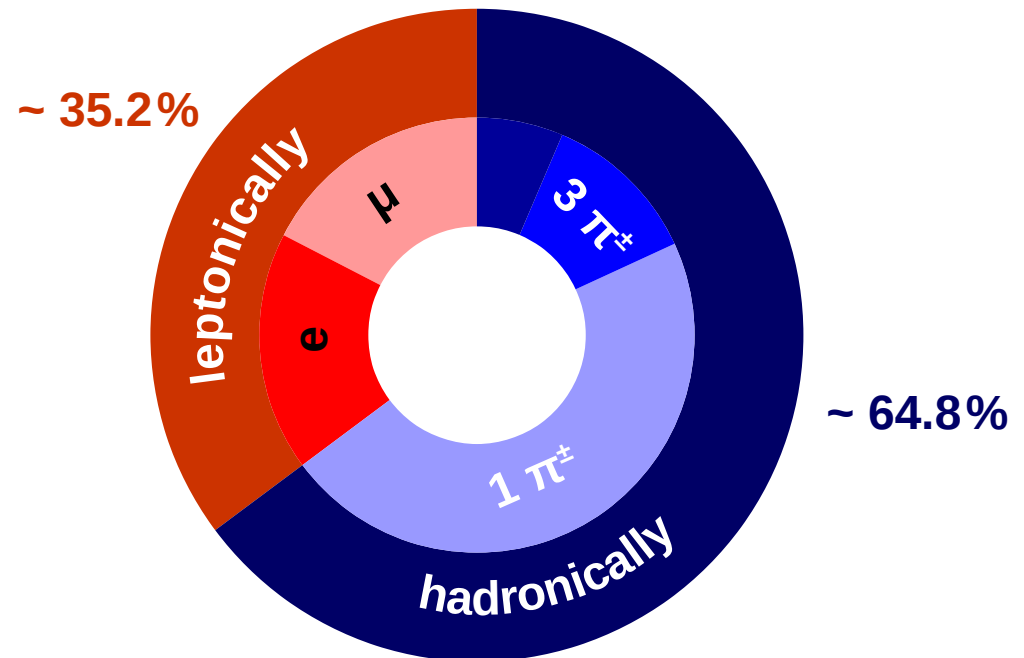
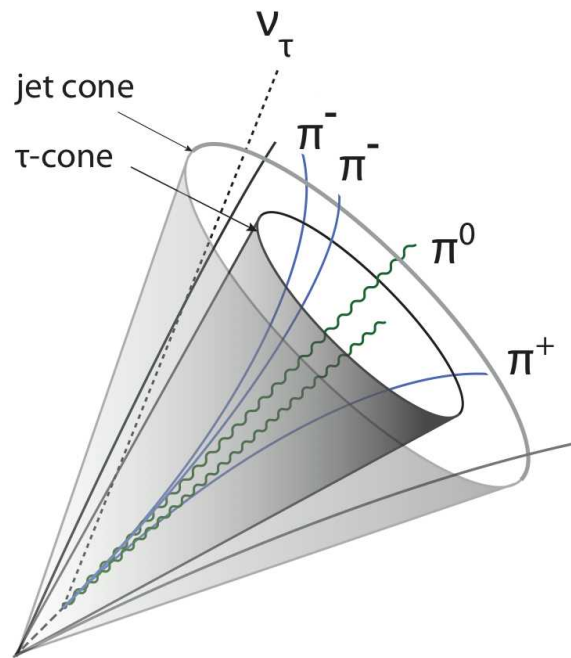
DPG Meeting, Würzburg
22.03.2018

Overview

- Tau reconstruction and identification important in many analyses
 - SM Higgs $\rightarrow \tau \tau$
 - BSM A/H/h/Z' $\rightarrow \tau \tau$
 - $W' \rightarrow \tau \nu$
 - ...
- Major parts of hadronic tau reconstruction and identification:
 - **track classification**
 - tau-identification
 - ...

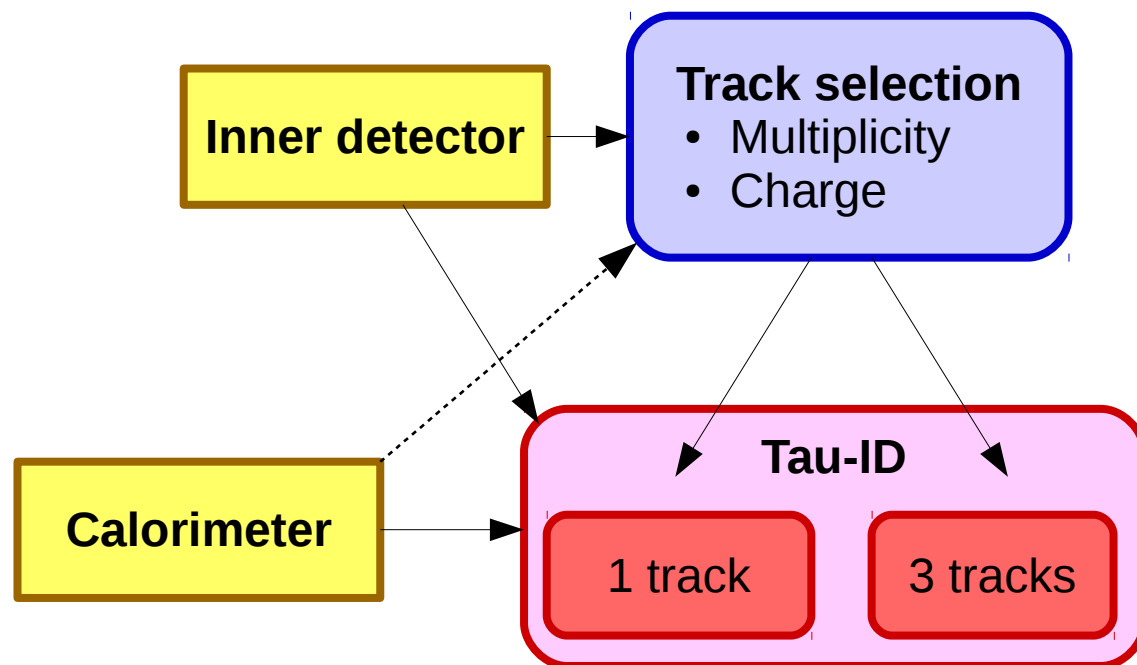
Reconstruction and identification

- Taus decay to a major part into hadrons like pions
- Hadronically decaying taus have odd charged track multiplicity
- Charged hadrons leave tracks in the inner detector of ATLAS



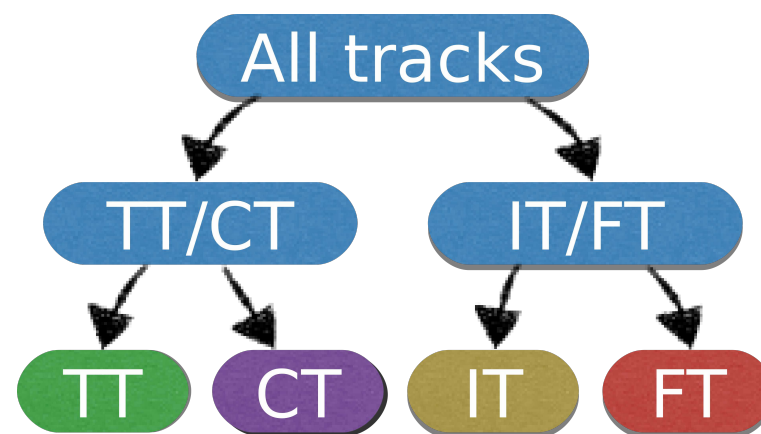
Reconstruction and identification

- **Task:** classify tracks and reconstruct charge of tau and charge multiplicity of decay
- Track selection run before Tau-ID
- Tau-ID algorithms depend on the charge multiplicity
- Build baseline for:
 - Tau-ID
 - PanTau
 - TES



BDT based track selection by Dirk Duschinger

- **Problem:** Classify tracks within the tau decay cone in multiple classes
 - **Tau tracks:** charged pions from tau decay
 - **Conversion tracks:** tracks from $\gamma \rightarrow e^+e^-$ conversions
 - **Isolation tracks:** underlying event
 - **Fake tracks:** everything else
- Currently done using multiple Boosted Decision Trees (BDT)
- Independent classification for each track

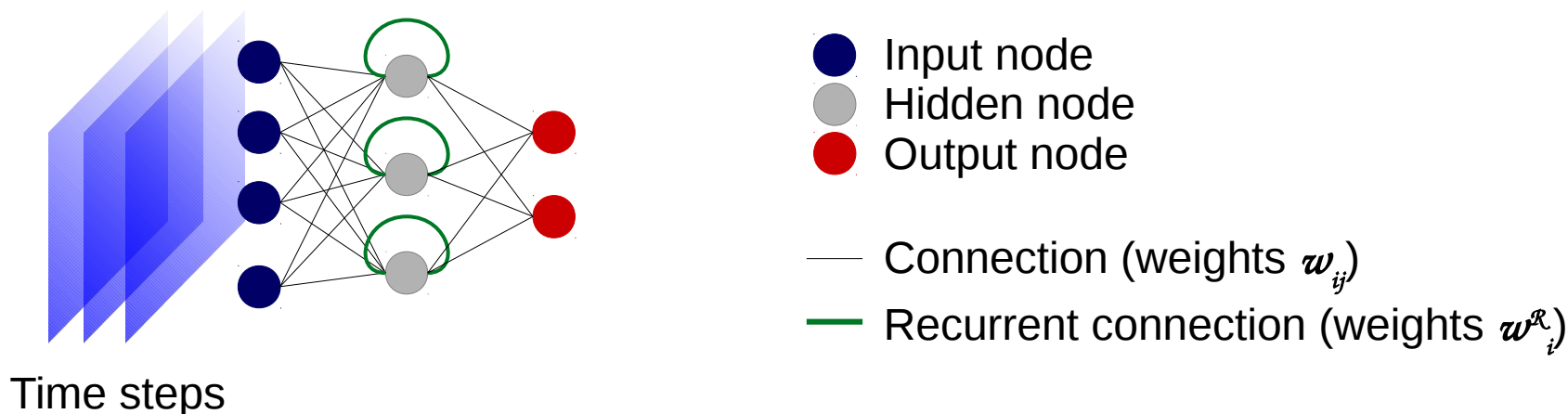


Recurrent neural networks

- **Problem:** Still residual information hidden in the correlation between tracks within one tau decay cone
- **Solution:** Classifier should take a look at all tracks and then make decisions
- BDT's can not easily handle input of variable length (~ 8-15 tracks per tau cone)
- Neural networks are more flexible in their use
- Recurrent neural networks allow making decision based on input variables **and** previous inputs

Recurrent neural networks

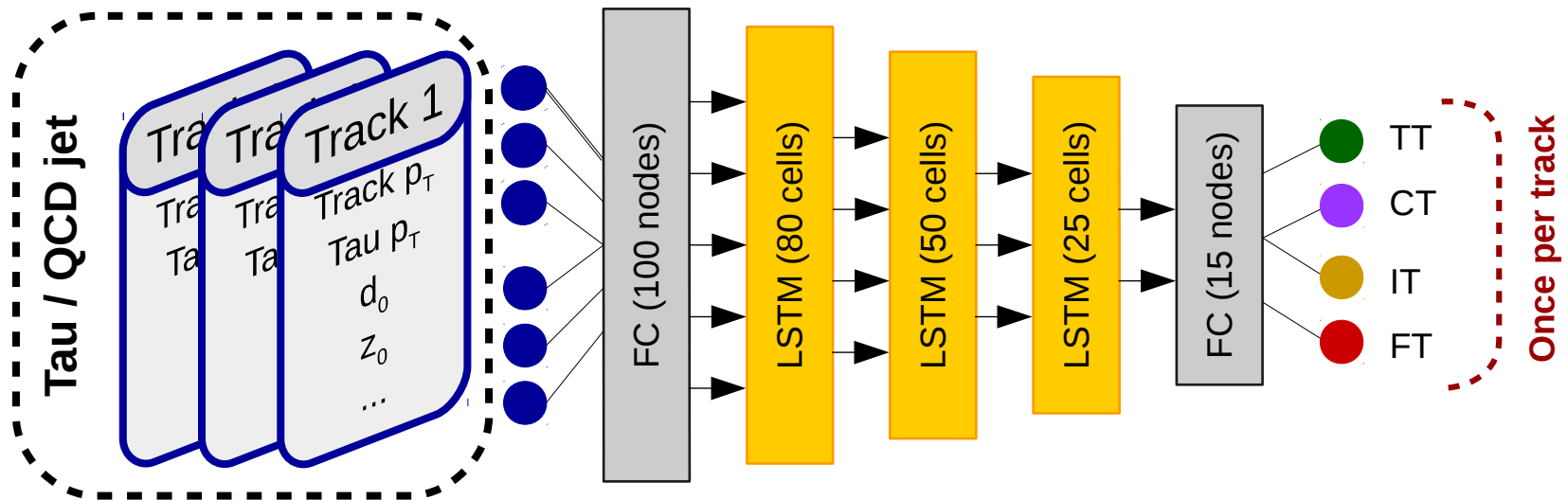
- Recurrent neural networks contain layers with nodes that reconnect to themselves and propagate decisions through time



- Input and output are **Fully Connected (FC)** layer
- Here actually **Long Short Term Memory (LSTM)** cells are used

Strategy

- RNN trained on taus and QCD jets \rightarrow minimize background in relevant kinematic regions



- Tracks ordered in p_T
- Between taus network state is reset
 \rightarrow **no correlation between individual taus**

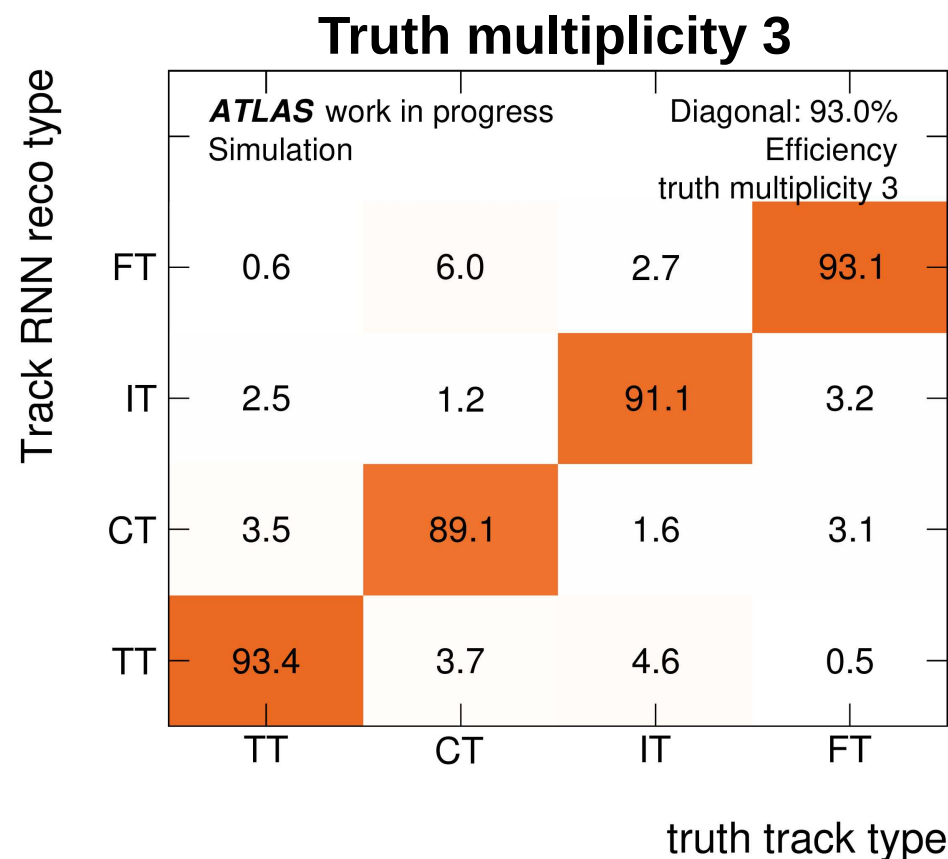
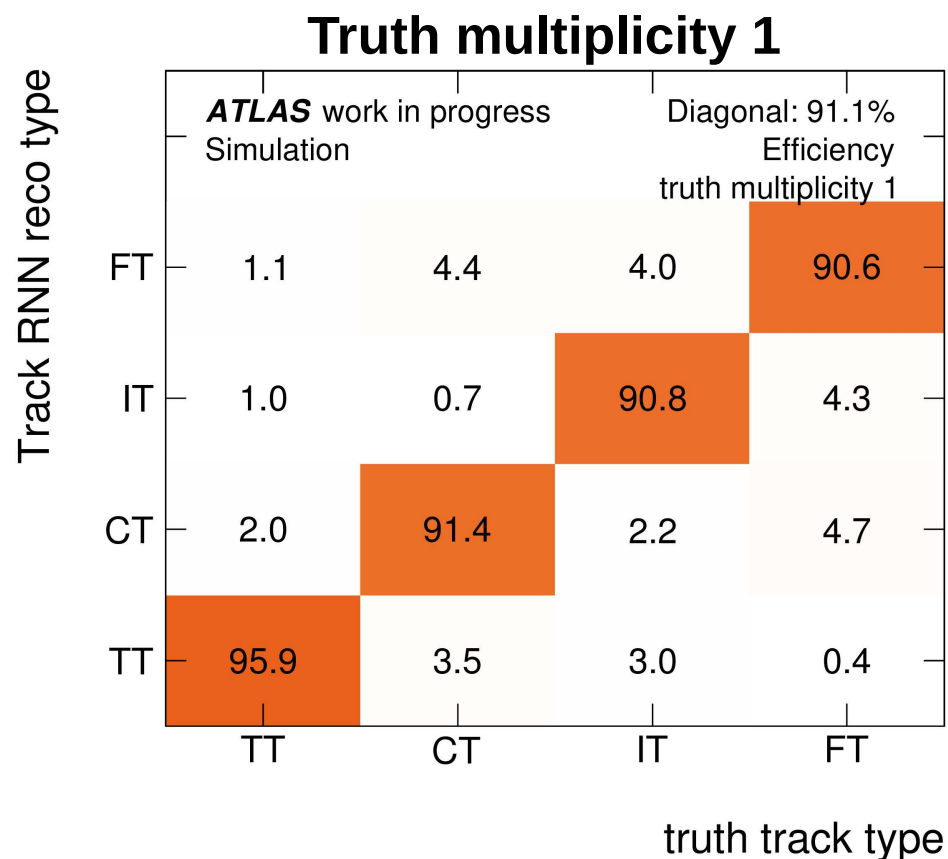
Results

TT : Tau Tracks

IT : Isolation Tracks

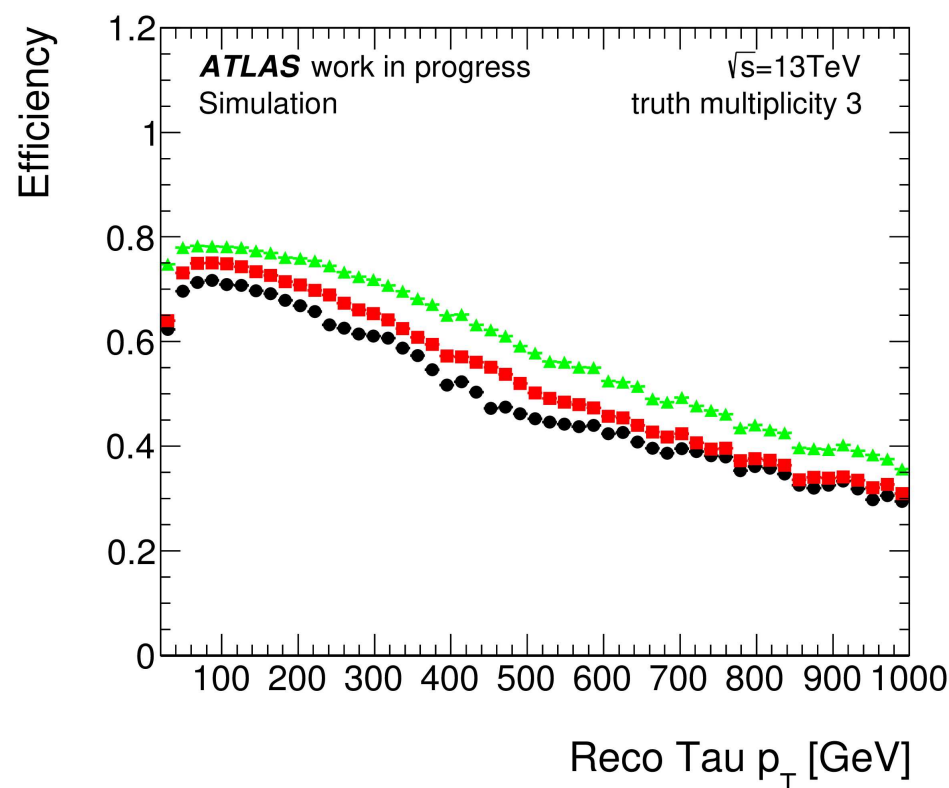
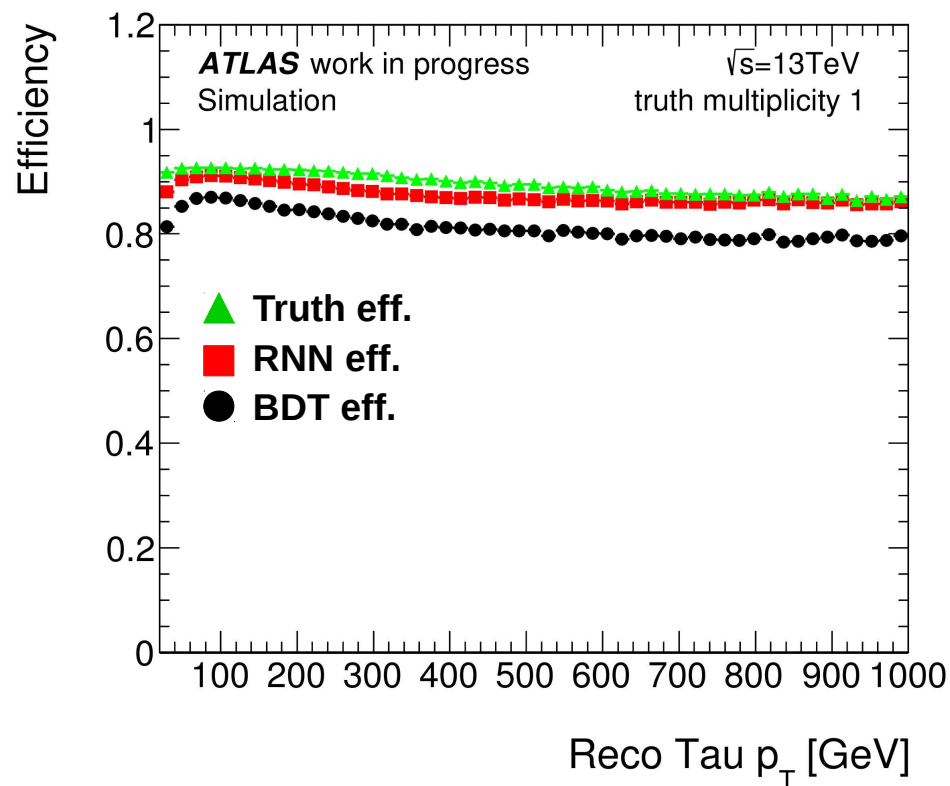
CT : Conversion Tracks

FT : Fake Tracks



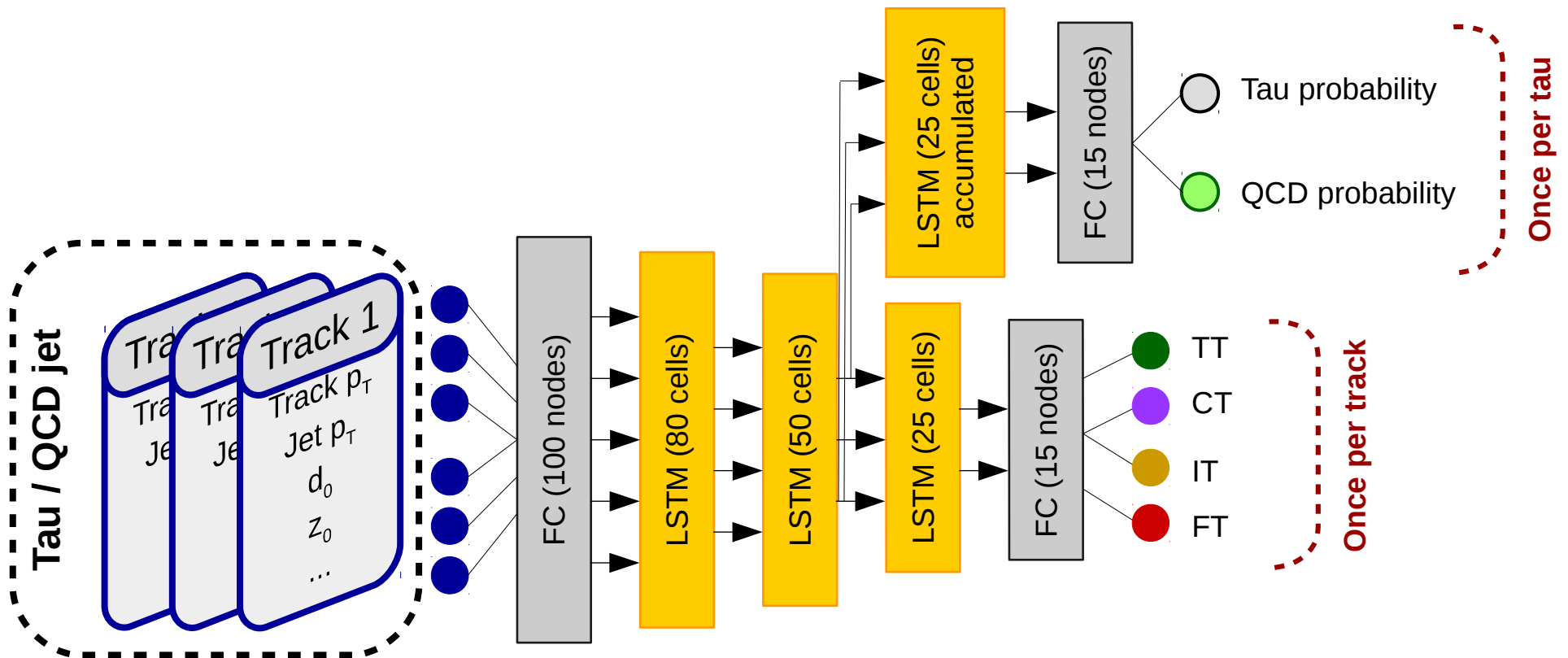
Results

- Reconstruction efficiency for taus with one charged track almost ideal



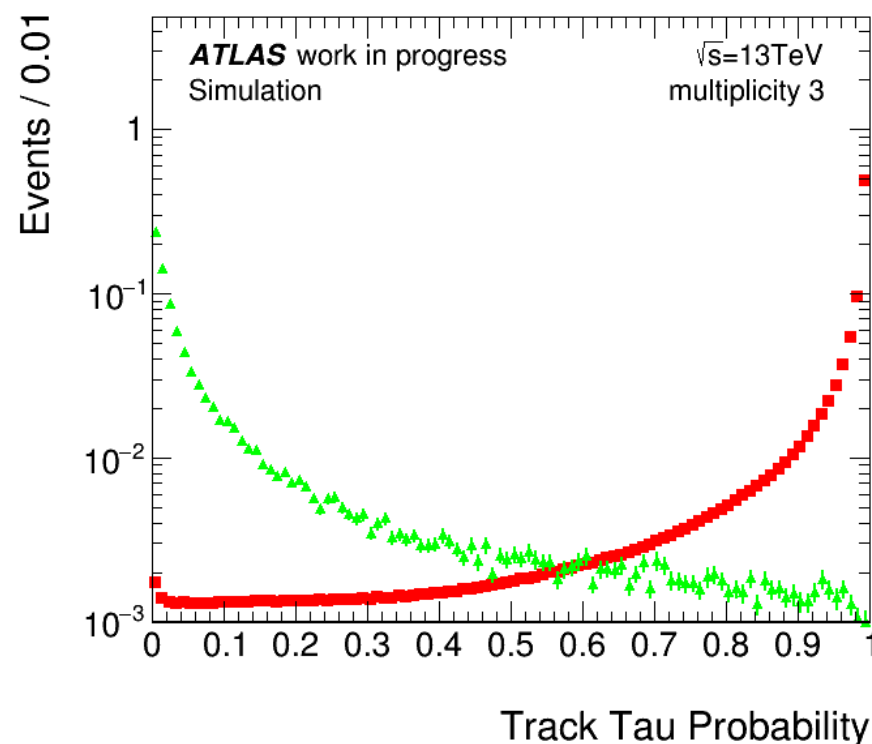
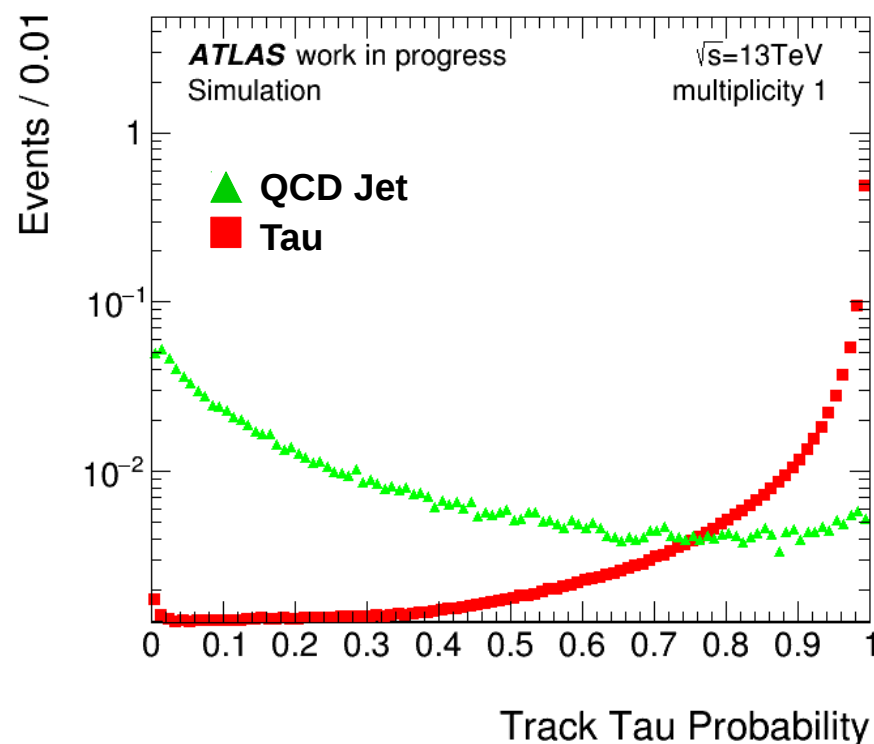
Strategy

- Additional branch for differentiating taus from QCD jets
- Calculates probability accumulated over all tracks



Results

- Additional tau probability produced based on tracks along the way
- Can be used as high level variable for Tau-ID



Conclusion

- RNN improves reconstruction efficiencies of tau decays with 1 charged track to almost maximum
- Reconstruction efficiencies for tau decays with 3 charged tracks increased

More taus for Tau-ID to separate against background

Additionally:

- New high level variable discriminating QCD jets against taus
- Might improve RNN based Tau-ID even more (see presentation by Christopher Deutsch)

Outlook

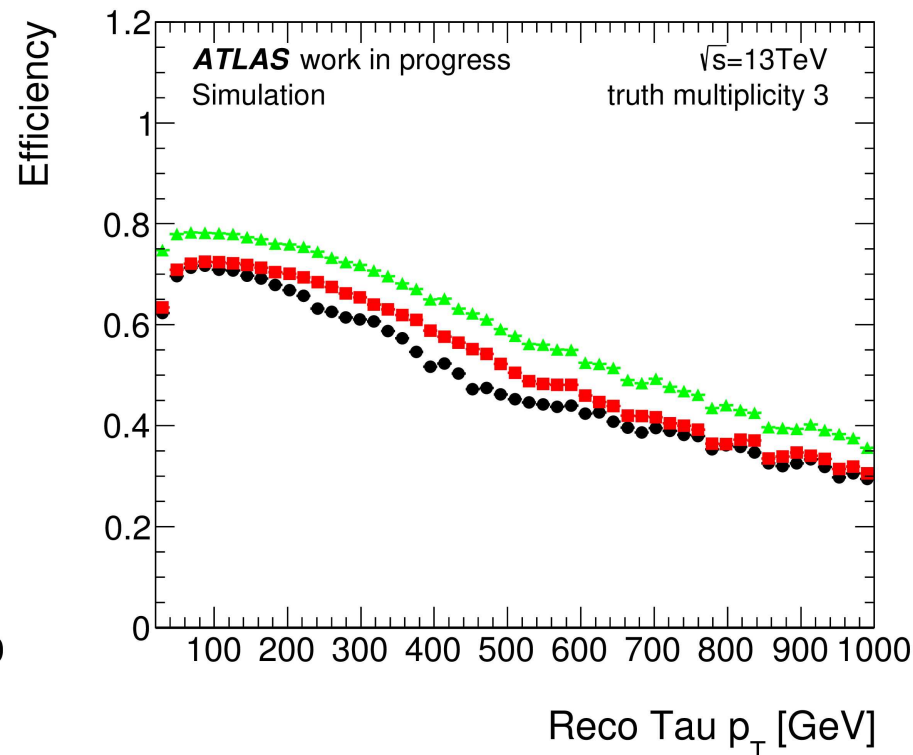
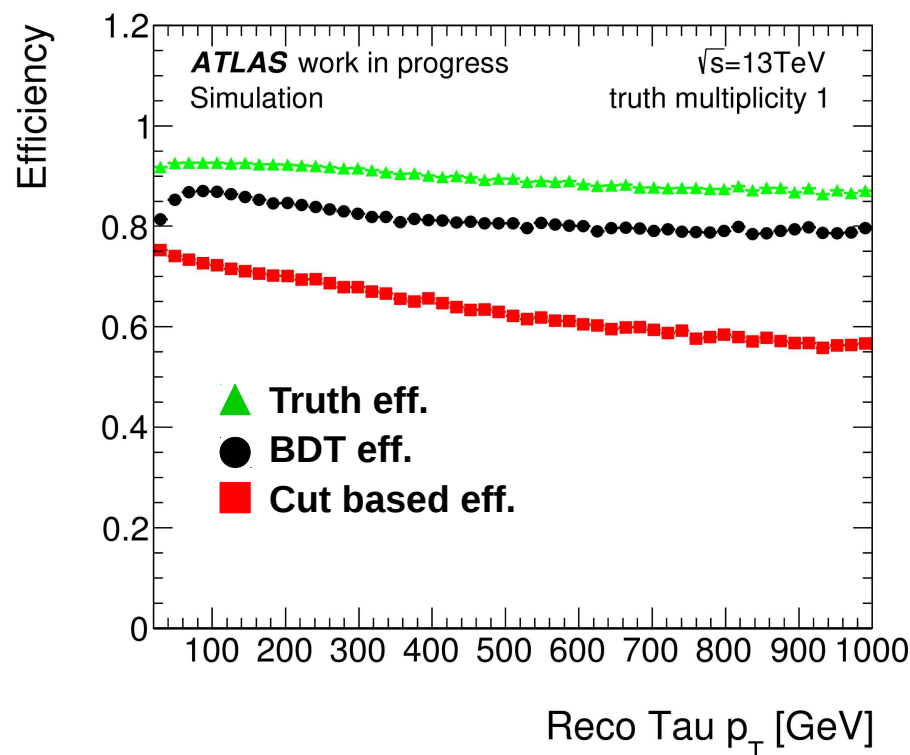
- New RNN based track classification has to be propagated through tau identification chain
- Tends to push more background in relevant signal regions → can Tau-ID mitigate this effect?
- Open tasks:
 - Study RNN based track selection + Tau-ID on QCD background
 - Study systematics

Still a long way to go but results seem promising

Backup

BDT based track selection

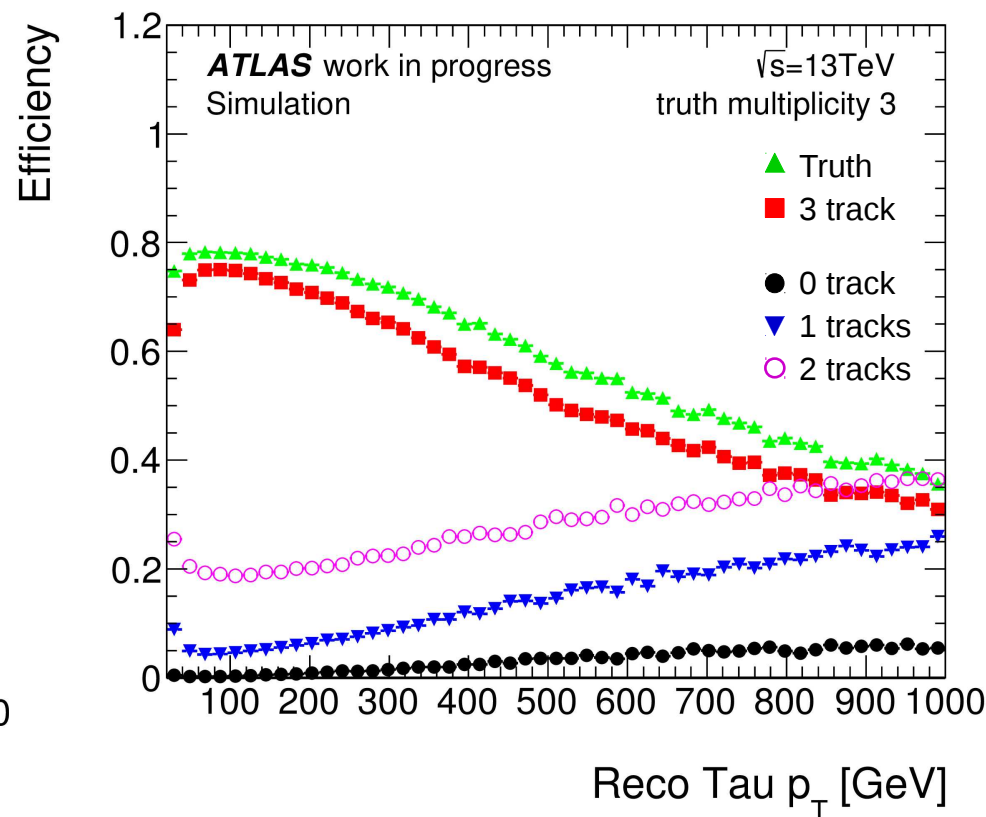
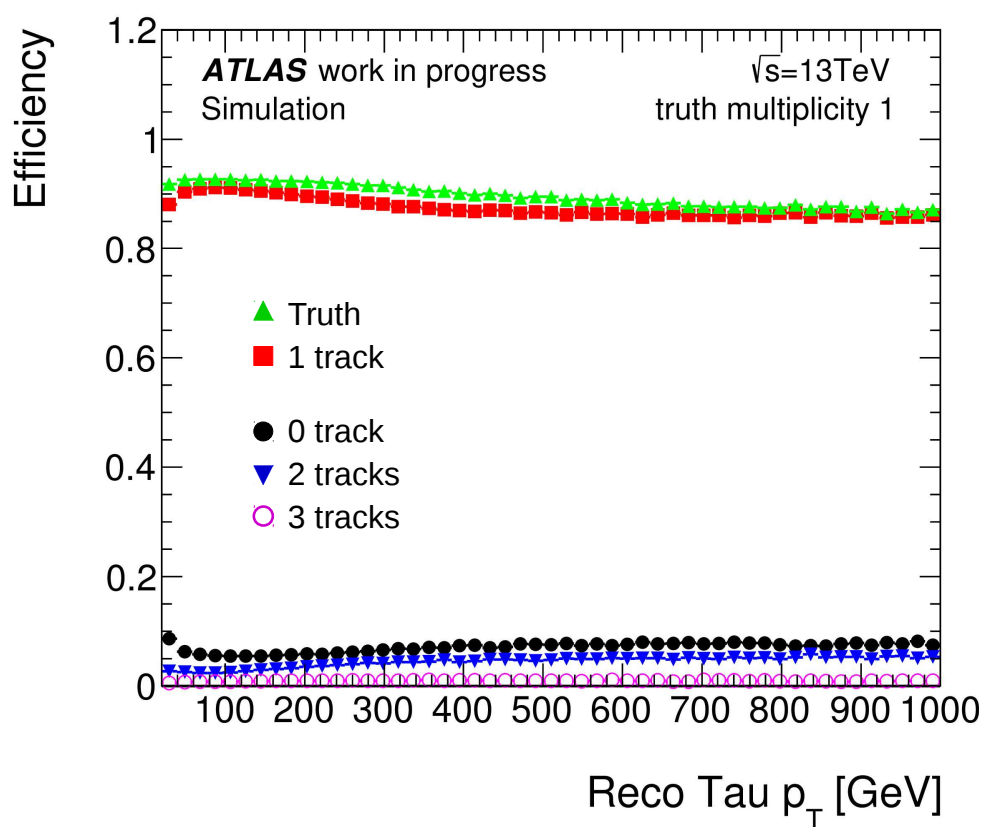
- BDT's Classifies each track independent from one another
- Show high reconstruction efficiency for tau decays with multiplicity 1 but worse for 3 than cut based



by Dirk Duschinger

Reconstruction inefficiencies

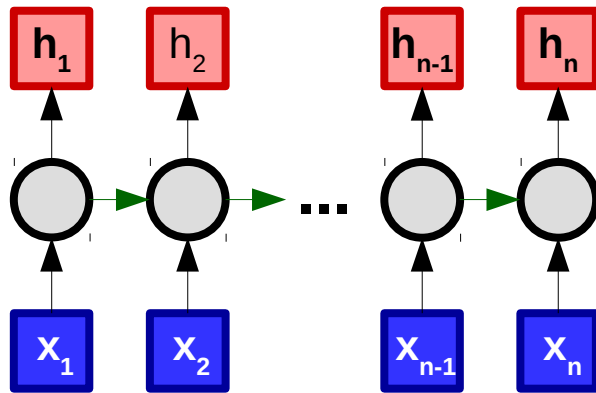
- High tracks of taus with p_T can not be resolved by track reco
- Tracks merge, e.g.: tau with 3 tracks \rightarrow 2 tracks



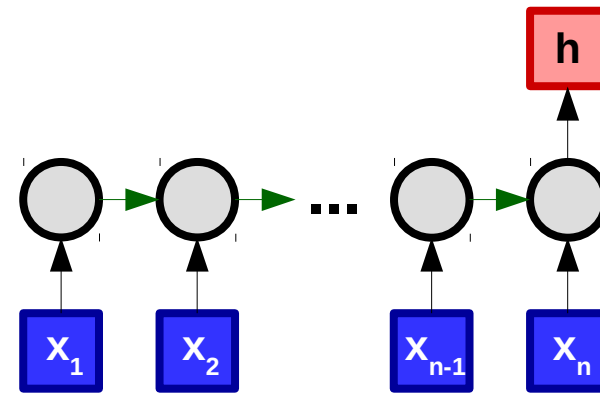
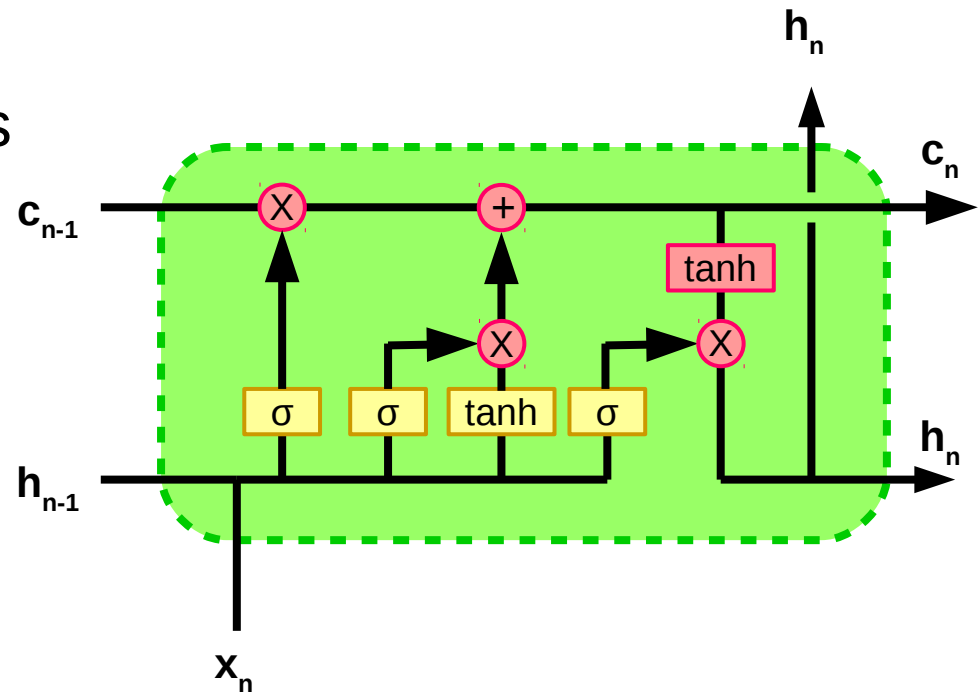
The LSTM cell

- Contains multiple internal nodes (gates i_t , o_t , f_t)
- Constructed to counteract vanishing gradient problem

Network unfolded in time:



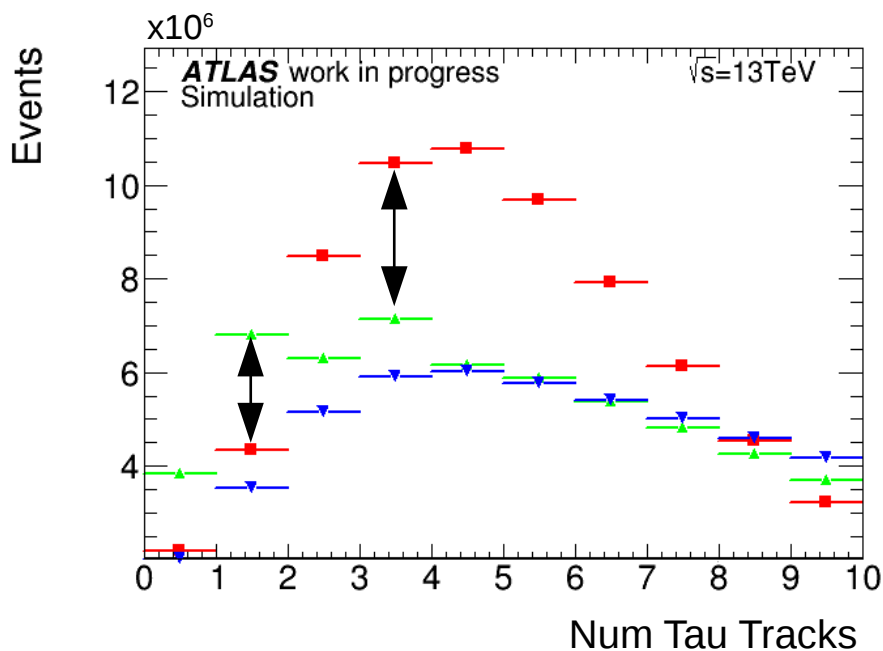
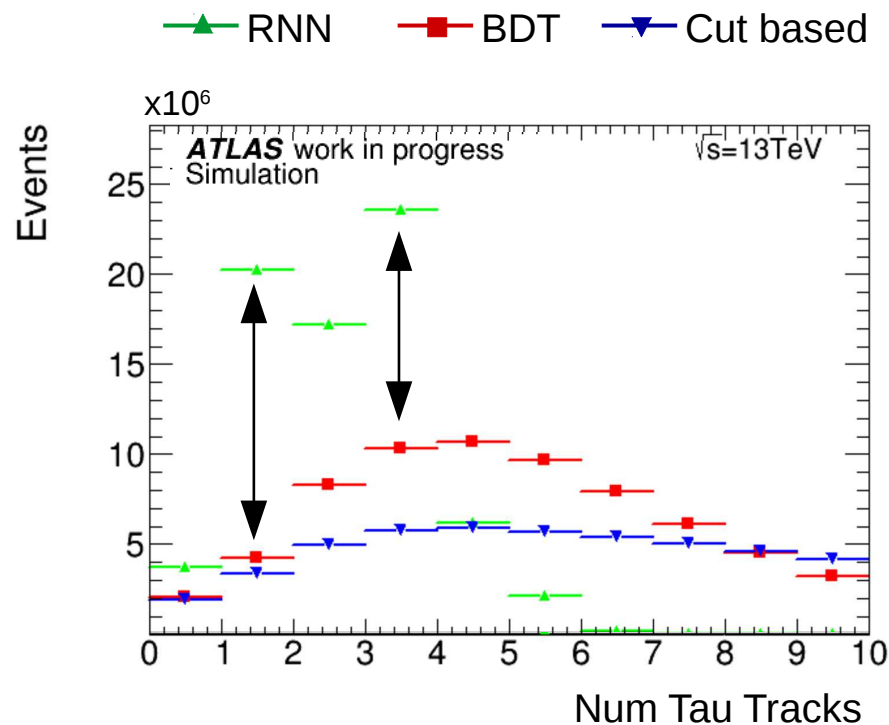
Many to many
(track selection)



Many to one
(Tau-QCD probability)

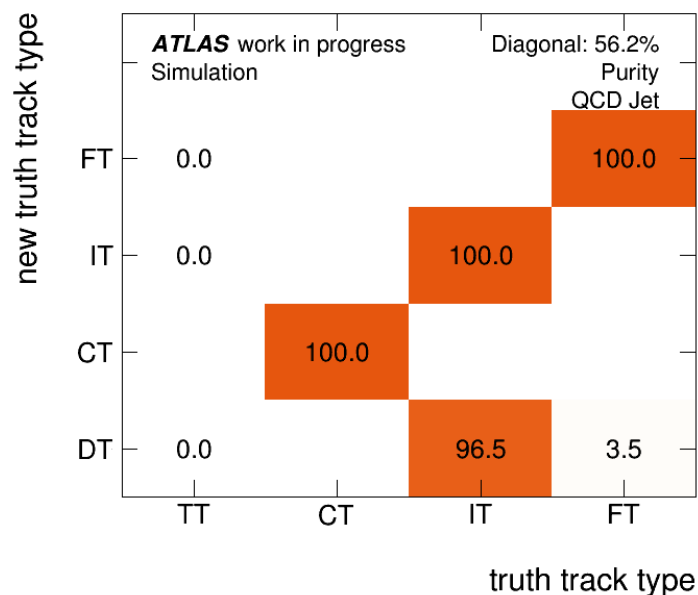
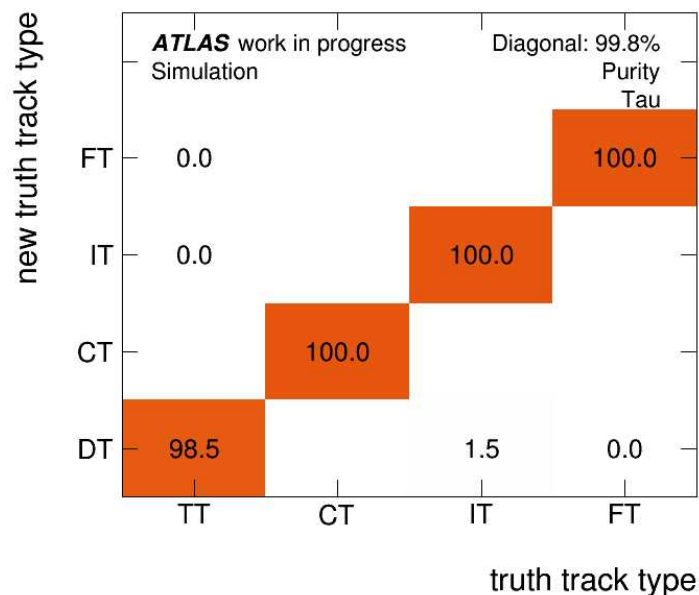
Effect of adding Background

- Background contamination in 1 and 3 tau track regions suppressed with QCD jets in training
- Even lower background in 3 track category



Effect of adding Background

- For training: redefinition **TT** → **Direct Track (DT)**
- DT: tracks fitted to first stable (by generator) generation of decay particles coming from highest energy tau/gluon/quark
- Like TT for taus, also works on jets
- **Problem:** generator dependent, use standard definition at runtime



Pile-up stability

- Stable over large range of μ

