Improved Selective Background Monte Carlo Simulation at Belle II with Graph Attention Networks and Weighted Events

Boyang Yu, Nikolai Hartmann, Luca Schinnerl, Thomas Kuhr

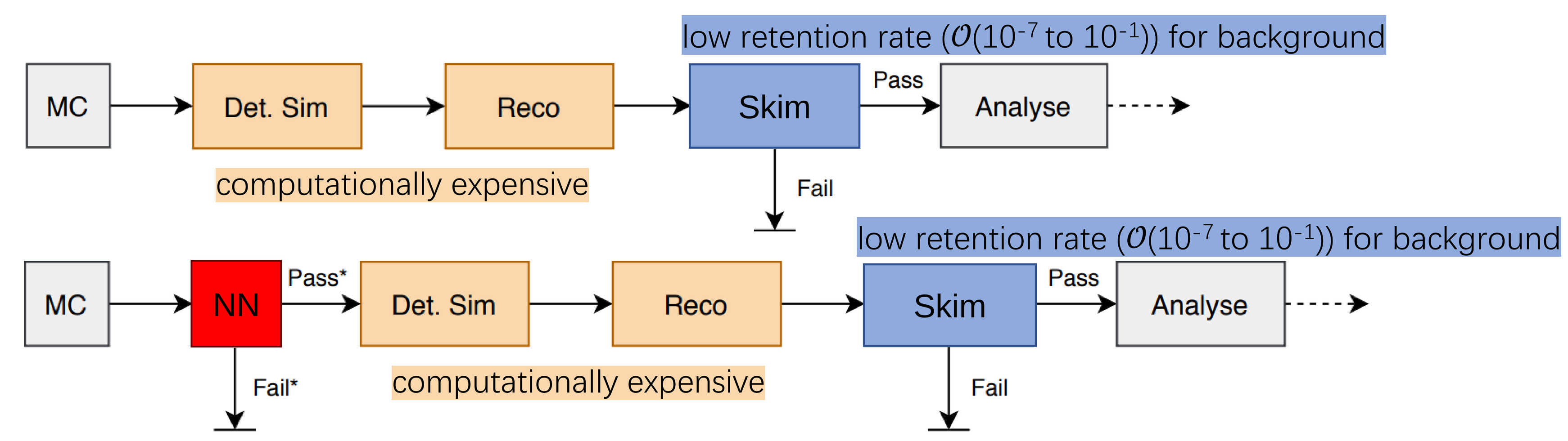
Contact: Boyang.Yu@physik.uni-muenchen.de

ACAT 2022 BARI

Belle II | LMU LUDWIG-MAXIMILIANS-UNIVERSITÄT MÜNCHEN | Bundesministerium für Bildung und Forschung | ORIGINS Excellence Cluster

Motivation:

- The measurements of rare processes require a huge luminosity. It means large simulations with low retention rates for background samples
- To avoid wasting of computation costs, filters for background using neural networks are introduced after the Monte Carlo event generation to select useful events for the following steps
- To deal with the bias introduced by filtering, different bias mitigating methods are studied

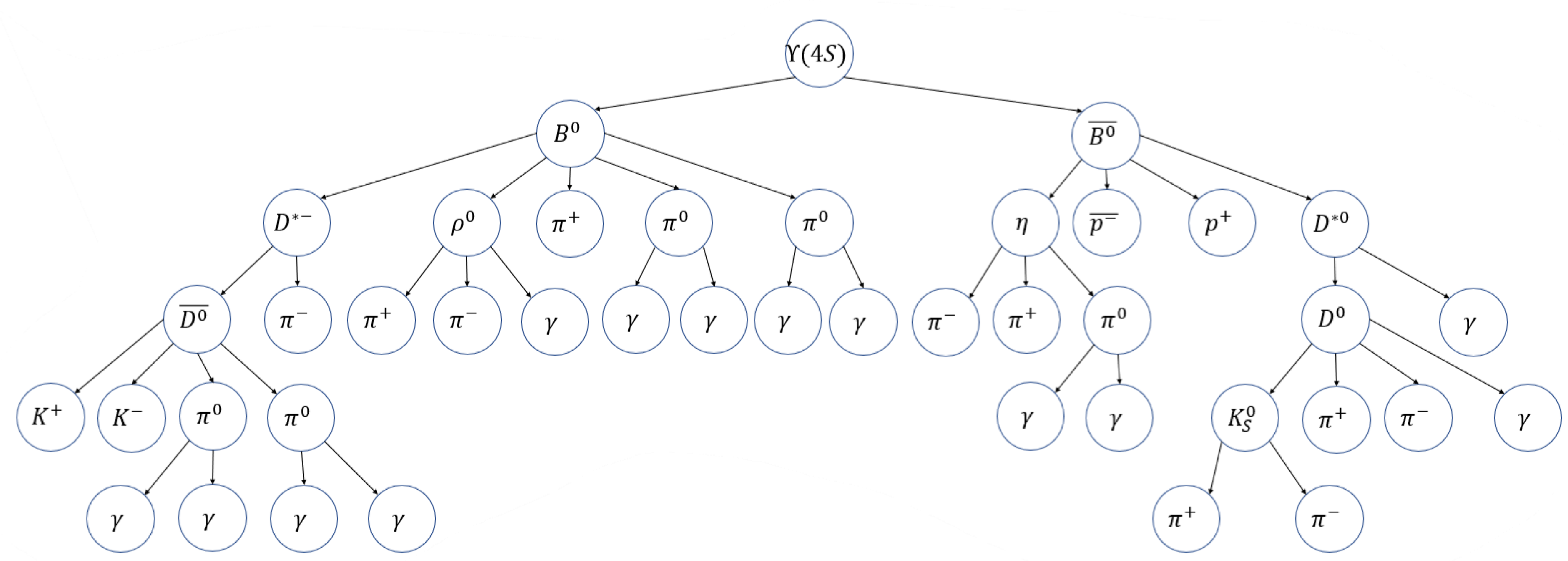


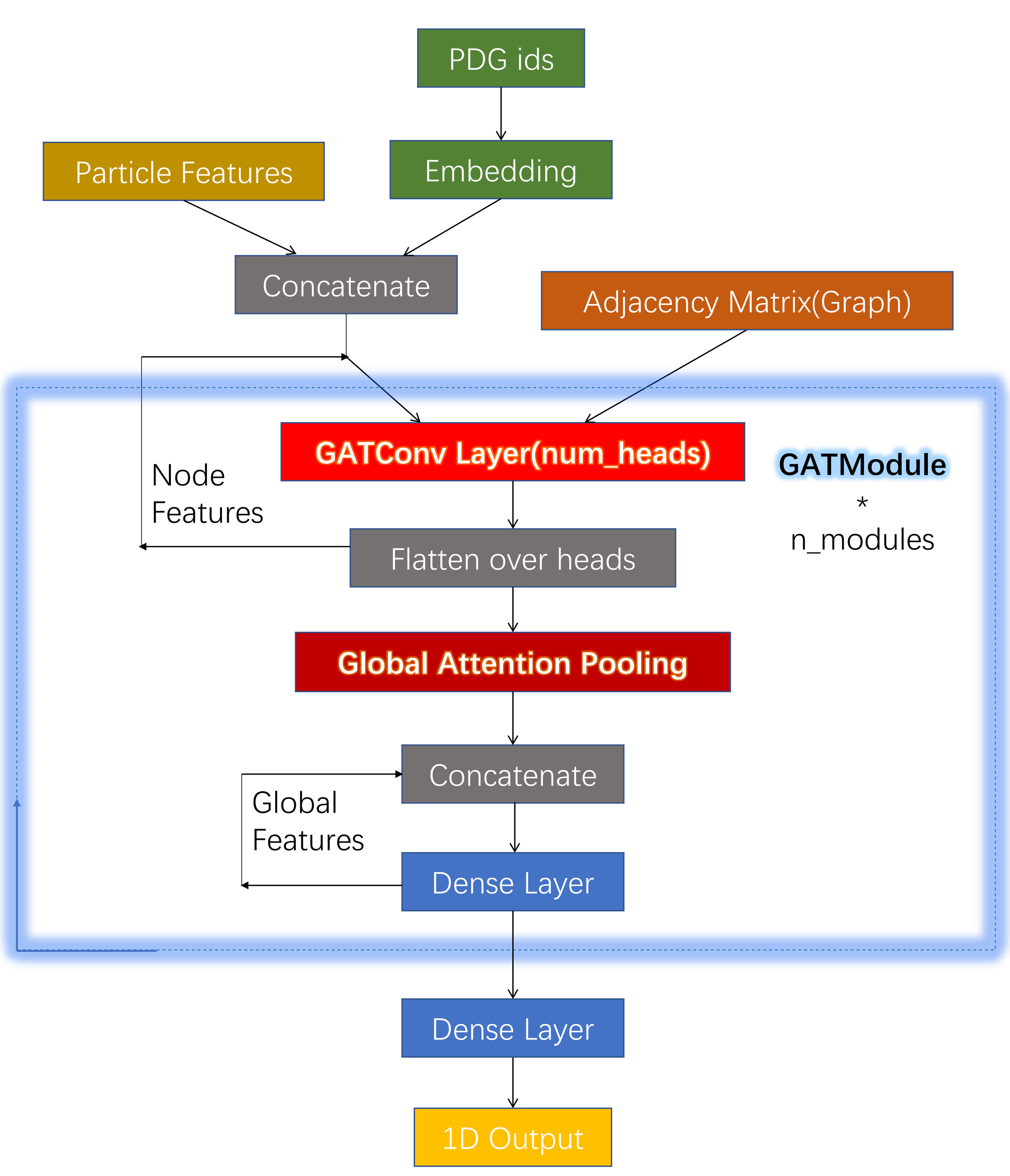
Previous works:

- Tree structure of particle decays -> Graph Convolutional Networks
- To quantify the bias, event-level and generator-level variables are introduced
- To mitigate the bias, a distance correlation loss is used in the training of the network

Dataset:

- Each event (each Graph):
  - Decay of $\Upsilon(4S) \to B^0\bar{B}^0$
  - Particles (Nodes)
  - Mother/Daughter relations (two way Edges) + self loops
  - Each particle (each Node)
    - PDG ids
    - 8 Features: Production time, Energy, Position (3d), Momentum (3d)
- Labels per event: Pass/Fail after the reconstruction of B decays (FEI skims)
- Other event level attributions for further analysis: e.g. M_{bc}, etc.





Improved filter:

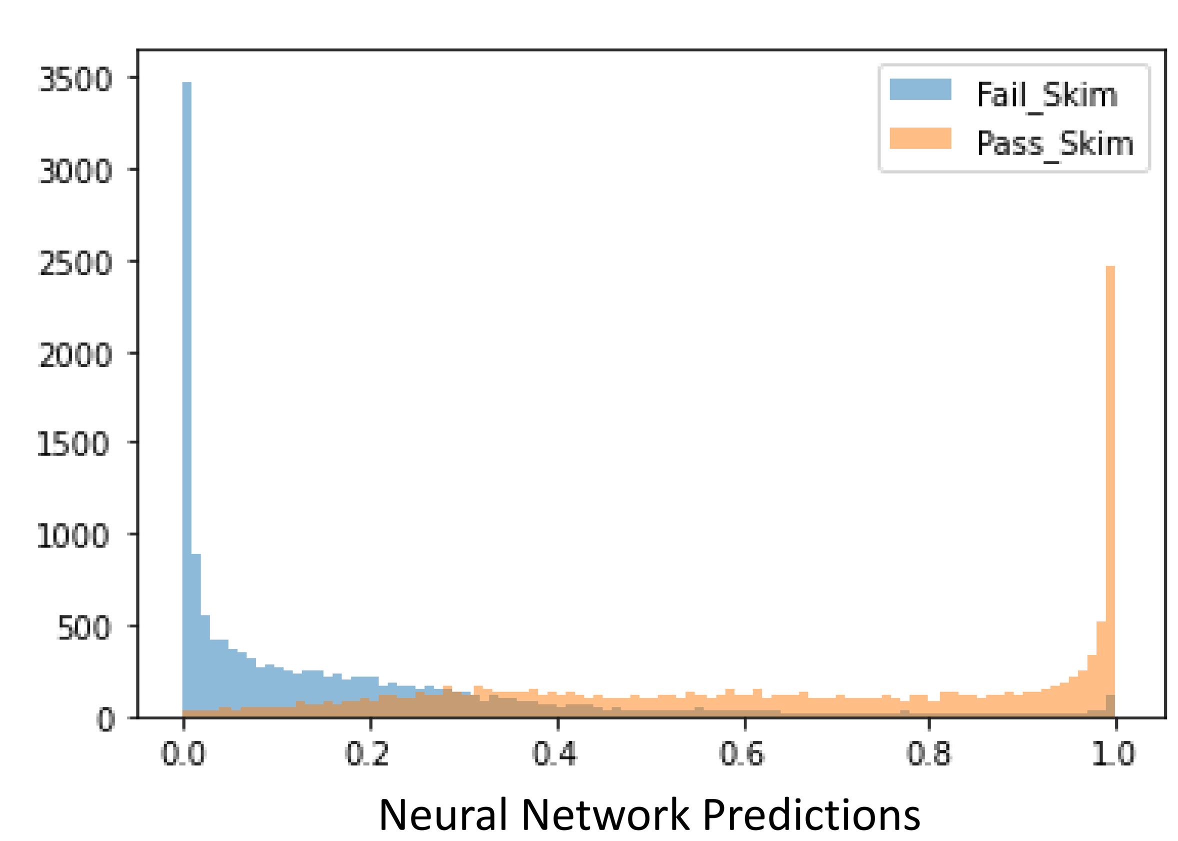
- Based on Graph Attention Networks instead of Graph Convolutional Networks
- Both node features for each particle and global features collecting all the information of nodes are updated simultaneously
- Global features are used in the final step to make predictions
- Best AUC value improved from 0.9083 to 0.9122

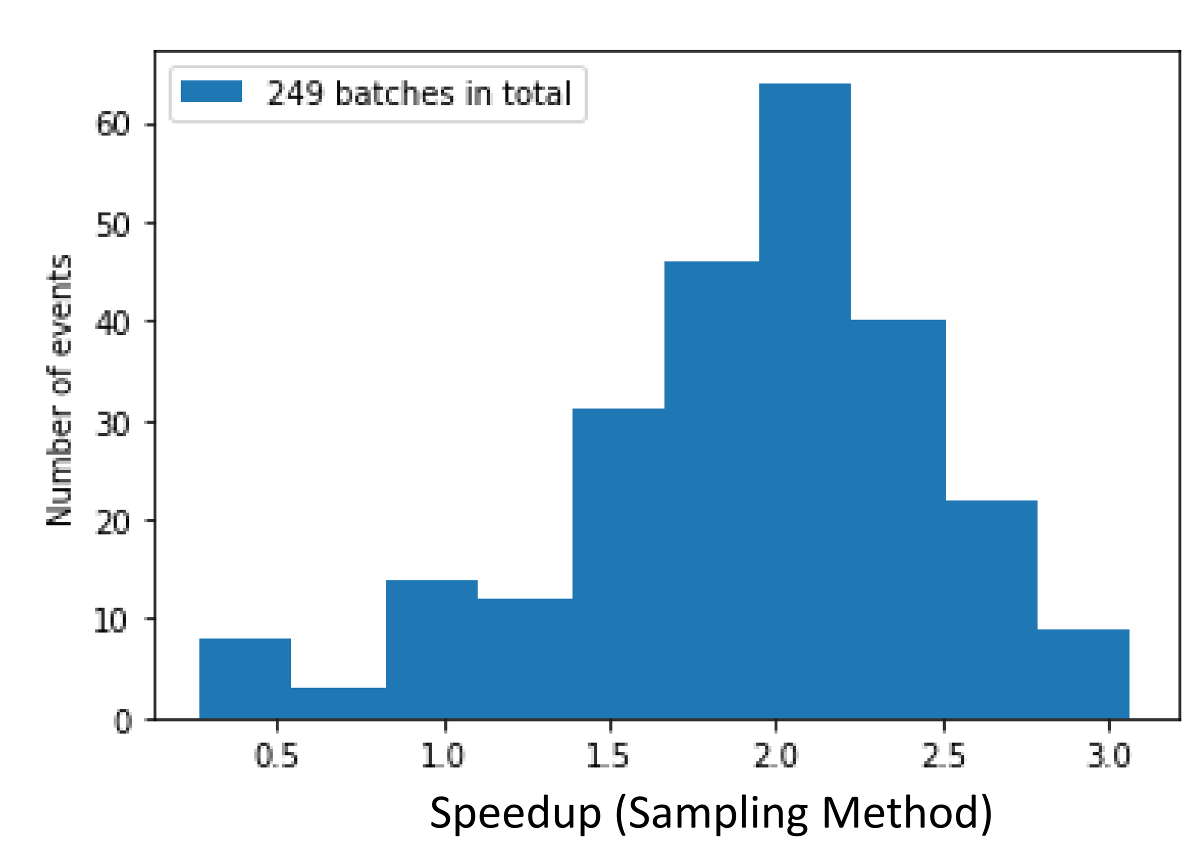
Speedup:

Improvement of computation time to produce the same effective number of events:

Speedup: $s = \frac{t_{no\_filter}}{t_{filter}}$

Effective Sample Size: $N_{eff} = \frac{(\sum \omega_i)^2}{\sum \omega_i^2}$





Improved bias reduction:

- Processed on the neural network outputs instead of adding loss functions in the training
- Studied two methods: Sampling and Reweighting
- Characteristics and performances are shown below

Bias mitigation method	Distance correlation loss	Reweighting	Sampling
Use of NN output	As score for selection according to fix threshold	As score for selection according to fix threshold	As probability to keep event randomly
Weight	No reweightings	Decided with the help of another classifier	Inverse of NN output
Loss to train NN	Binary cross entropy + distance correlation	Binary cross entropy	Speedup
Speedup	5.7	6.5	2.0
Bias	Small bias on some of the variables	Small bias on some of the variables, half of the previous work	No bias

Conclusion:

- Attention mechanism can improve NN performance for selective background monte carlo simulation
- Bias is avoided with sampling method while a speedup of factor 2 can still be maintained.
- Reweighting method can reach much higher speedup up to 6.5 but will still have some bias in the variables that are not used in the training of the extra classifier