Machine Learning for New Physics in $B \rightarrow K^* \mu^+ \mu^-$ Decays

Abstract
We report the status of a neural network regression model trained to extract new physics (NP) parameters in Monte Carlo (MC) data. We utilize a new EvtGen NP MC generator to generate $B \rightarrow K^* \mu^+ \mu^-$ events according to the deviation of the Wilson Coefficient $C_9$ from its SM value, $\delta C_9$, for different $\delta C_9$ values. We train a three-dimensional ResNet regression model, using images built from the the angular observables and the square of the invariant mass of the di-muon system, to extract values of $\delta C_9$ directly from MC data samples. This work is intended for future analyses at the Belle II experiment but may also find applicability at other experiments.

Decay Topology
Decay topology of a generic $B \rightarrow K^* \ell^+ \ell^-$ decay, showing the relevant angular observables used in neural network training.

Images
We produce “images” from generator-level MC, according to [1], that are used to train our neural network. Images are $\theta = M^2(\mu^+ \mu^-)$ values binned in bins of the angular observables. Our model is a three-dimensional, 34-layer, ResNet [2] trained to perform regression to extract Wilson Coefficient information, $\delta C_9 \equiv C_9^{\text{RES}} - C_9^{\text{SM}}$, directly from data[3].

Training History
From ensemble experiments, it is seen that the trained ResNet is able to correctly extract the different $\delta C_9$ values, from independent and unlabeled images. The black points are from experiments where the images are generated according to $\delta C_9$ values the ResNet has been trained with and the red points are from experiments where the images are generated according to $\delta C_9$ values with which the ResNet has not been trained.

Results

<table>
<thead>
<tr>
<th>( \chi^2 / \text{ndf} )</th>
<th>p0</th>
<th>p1</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.169 / 31</td>
<td>-0.02821 ± 0.00783</td>
<td>0.9778 ± 0.02596</td>
</tr>
</tbody>
</table>

From ensemble experiments, it is seen that the trained ResNet is able to correctly extract the different $\delta C_9$ values, from independent and unlabeled images. The black points are from experiments where the images are generated according to $\delta C_9$ values the ResNet has been trained with and the red points are from experiments where the images are generated according to $\delta C_9$ values with which the ResNet has not been trained.

References
[3] Done in collaboration with the authors of [1]